



# Contrainte des modèles génétiques de réservoirs par une approche de reconnaissance statistique de forme

Tatiana Chugunova

## ► To cite this version:

Tatiana Chugunova. Contrainte des modèles génétiques de réservoirs par une approche de reconnaissance statistique de forme. Mathématiques [math]. École Nationale Supérieure des Mines de Paris, 2008. Français. NNT : 2008ENMP1554 . pastel-00004452

**HAL Id: pastel-00004452**

**<https://pastel.archives-ouvertes.fr/pastel-00004452>**

Submitted on 12 Dec 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Président  
Rapporteur  
Rapporteur  
Examineur  
Examineur  
Examineur



## Remerciements

Avant confronter le lecteur au sujet principal de ce manuscrit, je me permets d'aborder un sujet moins scientifique mais néanmoins important : je voudrais remercier ici tous ceux et celles qui m'ont accueillie, ont aidée à m'installer et m'ont soutenue tout au long de mon séjour professionnelle en France, et plus particulièrement, au sein de l'IFP.

En premier lieu, je veux remercier mon promoteur de thèse, Lin Ying Hu, qui m'a exprimé sa confiance dès le début de la thèse. En me soutenant tout en me laissant une grande liberté d'action, Lin Ying m'a permis non seulement d'entreprendre ce travail de thèse mais aussi de satisfaire ma curiosité dans le domaine de la géostatistique (ce qui ne correspondait pas vraiment à ma formation de base). Je veux également remercier mon directeur de thèse, Christian Lantuéjoul, pour la précision de ses explications, sa patience, sa capacité à traduire mes pensées et les rendre compréhensibles. Je le remercie plus particulièrement pour le travail qu'il a accompli sur ce manuscrit, pour les millions (je me permets ici d'être statistiquement incorrecte !) de fautes corrigées, de phrases remaniées et de paragraphes restructurés.

En seconde lieu, je souhaite remercier tous mes collaborateurs de l'IFP : mon premier superviseur Olivier Vincké, mes camarades et mes collègues de la division : Ludovic Ricard, Alexandre Fornel, Carine Famy, Celine Scheidt, Hélène Cartier, Sylvie Hoguet. Je tiens tout particulièrement à remercier mes collègues de bureau et de l'étage de me supporter telle que je suis et de rendre les journées de travail plus conviviales : Magnolia Mamaghani, Alexandre Abellan, Charles Cohen, Amir Soltani, Samir Touzani. Merci aussi à mon collègue et compatriote russe, Pavel Spesivtcev, pour les déjeuners et les pauses thé partagées ensemble ainsi que pour m'avoir aidée à surmonter divers problèmes liés à la vie à l'étranger. Je tiens à remercier Virginie Patacz, qui a eu la gentillesse de travailler sur le texte brut afin d'éliminer des fautes d'orthographe et de grammaire et d'éclaircir nombre d'idées que j'ai voulu faire passer.

Je veux remercier ma hiérarchie : Vincent Richard, Patrick Lemonnier et Frederic Roggero, qui ont souvent fermé les yeux sur mes arrivées tardives mais qui n'ont pas manqué, malgré tout, d'apprécier mon travail. Un remerciement spécial va à Benoît Noetinger, qui a suivi de près l'avancement de ma thèse et qui m'a toujours apporté son soutien lors de problèmes administratifs. Grâce à son aide efficace, j'ai toujours le plaisir d'être en France et de proposer ce manuscrit à votre attention.

Je n'oublierai pas la contribution essentielle de mes professeurs de Moscou à mon parcours scientifique. Non seulement ils m'ont appris différentes disciplines scientifiques, mais ils m'ont aussi transmis une partie de leur curiosité, la soif de savoir et l'envie de comprendre. Merci à Elena Valerievna Glivenko, Vladimir Markovich Entov, Mihail Georgievich Suharev, Sergey Sergeevich Arseniev-Obrazcov, Nikolai Iosipovich Osetinskiy. Низкий вам поклон.

Merci à mes parents qui m'ont toujours soutenu dans tous mes projets sans me rapprocher de m'éloigner d'eux géographiquement ni douter de mes décisions. Спасибо вам, мои дорогие мама и папа за то, что всегда поддерживаете меня во всех моих начинаниях; за то, что не упрекаете в том, что я от вас далеко; за то, что не сомневаетесь ни во мне, ни в решениях, которые я принимаю.

## Résumé :

L'objet de cette thèse est la simulation de la structure interne de réservoirs pétroliers hétérogènes. Pour plus de réalisme, nous avons recours à des modèles génétiques basés sur des processus sédimentaires. Les simulations obtenues contiennent des structures géométriques complexes (réseaux de fractures, chenaux méandriformes ....), et sont bien plus à même de représenter un réservoir souterrain que celles produites par les modèles stochastiques traditionnels. Il reste à intégrer toute l'information pertinente auxiliaire (sismique, diagrapie, mesures des carottes), ce qui demeure une tâche difficile.

Pour ce faire, nous avons adapté les techniques de simulations multipoints (MP).

Dans son principe, l'approche MP classique est un algorithme séquentiel. A chaque nouveau point à simuler est associée une configuration de valeurs prises en des points de son voisinage. On repère dans une image d'apprentissage les positions où figure cette configuration. On forme ensuite l'histogramme des valeurs en ces différentes positions. Finalement on affecte au nouveau point une valeur simulée selon cet histogramme. Une étude approfondie nous a permis de préciser les forces, les faiblesses ainsi que les limites d'application de cette approche.

L'inconvénient majeur de l'approche MP est qu'elle n'est applicable que moyennant une certaine stationnarité spatiale du réservoir. Pour étendre cette approche à la simulation de réservoirs non stationnaire, deux algorithmes ont paru dans la littérature. Le premier algorithme, dit modèle Tau, propose une modélisation paramétrique d'un histogramme conditionnel. Malheureusement, cet algorithme ne permet de prendre en compte qu'un seul type de contraintes spatiales (e.g. des proportions locales de faciès). Son paramètre est en général difficile à inférer, et les simulations qu'il produit révèlent des déformations géométriques importantes. Le deuxième algorithme, dit de classement, part de deux images d'apprentissage (l'une pour l'information principale, l'autre pour l'information auxiliaire). Lors de sa mise en œuvre, il suppose les contraintes spatiales discrétisées.

Dans ce travail, nous avons proposé un autre algorithme pour intégrer des contraintes auxiliaires aux simulations MP dans un contexte non stationnaire. Cet algorithme n'est pas bâti sur un modèle probabiliste particulier (contrairement au modèle Tau), et il permet d'intégrer effectivement des contraintes spatiales continues (ce qui n'est pas le cas de l'algorithme de classement). Cet algorithme a été comparé aux algorithmes existants sur deux cas synthétiques (fonction aléatoire gaussienne seuillée, schéma booléen d'objets) et sur un cas réel (réseau de fractures). Les résultats fournis par notre algorithme sont très encourageants. Dans tous les cas, les contraintes spatiales sont bien restituées, aussi bien globalement que localement, ce que ne parviennent pas à faire le modèle Tau et l'algorithme de classement.

Cet algorithme a aussi été appliqué sur un système de chenaux méandriformes simulé par une approche génétique. Les résultats ont été comparés avec ceux obtenus par l'approche MP classique. En fin de compte, les résultats expérimentaux obtenus confirment que notre algorithme a un domaine d'applicabilité bien plus vaste que les approches MP classique et étendues. Il permet de reproduire la complexité de la géométrie multifaciès de réservoirs pétroliers, même de forts trends spatiaux.

**Mots clés :** approche statistique multipoints, simulation stochastique non stationnaire, contrainte spatiale, modélisation génétique.





# Table des Matières

<b>1</b>	<b>Introduction</b>	<b>11</b>
1.1	Réservoir pétrolier.....	12
1.2	Modélisation numérique d'un réservoir.....	12
1.3	Modélisation géologique.....	15
1.3.1	Approche génétique.....	15
1.3.2	Approche stochastique.....	16
1.3.3	Approche statistique multipoints.....	19
1.4	Cadre de la thèse.....	21
<b>2</b>	<b>Approche statistique multipoints</b>	<b>23</b>
2.1	Présentation générale.....	23
2.1.1	Principe.....	23
2.1.2	Algorithme initial.....	25
2.2	Mise en oeuvre.....	28
2.2.1	Scanning de l'image d'apprentissage.....	30
2.2.2	Simulation.....	31
2.3	Conditions d'application.....	33
2.3.1	Taille de l'image d'apprentissage.....	33
2.3.2	Fenêtre d'exploration.....	34
2.3.3	Chemin de simulation.....	35
2.3.4	Connexité.....	39
2.3.5	Stationnarité.....	47
2.3.6	Résumé.....	49
2.4	Vers la simulation non stationnaire.....	51
<b>3</b>	<b>Simulation non stationnaire : approches existantes</b>	<b>53</b>
3.1	Description du problème.....	54
3.2	Méthode de classement.....	55
3.2.1	Principe.....	55
3.2.2	Application.....	56
3.2.3	Remarques.....	57
3.3	Modèle Tau.....	58
3.3.1	Principe.....	58
3.3.2	Validation sur trois modèles stochastiques.....	60
3.3.3	Application.....	68
3.3.4	Remarques.....	74

<b>4</b>	<b>Simulation non stationnaire : l'approche MPNST</b>	<b>75</b>
4.1	Présentation Générale .....	75
4.1.1	Principe.....	75
4.1.2	Algorithme.....	77
4.1.3	Critère de compatibilité.....	78
4.2	Mise en oeuvre.....	80
4.2.1	Scanning de l'image d'apprentissage.....	82
4.2.2	Simulation .....	83
4.3	Exemples de validation.....	84
4.3.1	Fonction gaussienne seuillée.....	84
4.3.2	Réseau de fractures.....	86
4.3.3	Segments orientés.....	88
4.4	Résumé.....	90
<b>5</b>	<b>Application au modèle génétique</b>	<b>91</b>
5.1	Modèle génétique.....	91
5.2	Simulation par l'approche MP.....	94
5.3	Simulation par l'approche MPNST .....	97
5.4	Résumé.....	105
<b>6</b>	<b>Conclusions et perspectives</b>	<b>107</b>
6.1	Approches MP.....	107
6.2	Approches existantes .....	108
6.3	Approche MPNST.....	109
6.4	Application au modèle génétique .....	110
6.5	Perspectives.....	110
<b>7</b>	<b>Bibliographie</b>	<b>112</b>
	<b>Annexe A. Notation</b>	<b>120</b>
	<b>Annexe B. Caractérisation d'image</b>	<b>123</b>
B.1.	Statistiques monovariables .....	123
B.1.1.	Moyenne globale .....	123
B.1.2.	Moyenne locale.....	124
B.2.	Statistiques bivariables.....	125
B.2.1.	Variogramme .....	125
B.3.	Statistiques multivariables .....	126

Table des Matières	10
B.3.1. Connexité .....	126
B.3.2. Connexité multipoints .....	126
B.3.3. Statistiques multipoints .....	127
 <b>Annexe C. Critère progressif de compatibilité</b>	 <b>128</b>
C.1. Poids approximés par le modèle gaussien .....	128
C.2. Poids approximés par l'histogramme .....	129
C.3. Poids approximés localement .....	129

# 1 Introduction

Pour rendre cette thèse accessible au lecteur, et lui donner une vision globale de son contenu, nous voudrions en exposer la structure, en dégager les points clés et donner un aperçu de notre travail en fonction de son évolution.

Nous commençons par un chapitre d'**introduction**, qui porte sur l'ingénierie de réservoir et ses besoins. Après avoir brièvement présenté l'ensemble des étapes nécessaires à la modélisation d'un gisement pétrolier, nous allons nous concentrer plus particulièrement sur l'étape de modélisation géostatistique, ce qui constitue le cadre de cette thèse.

Le chapitre 2 est consacré à la description de l'**approche multipoints**. Une étude bibliographique est proposée, suivie d'une discussion détaillée de ses limites d'utilisation, tout particulièrement au niveau de la stationnarité. C'est précisément pour mieux restituer des structures géométriques non stationnaires que nous sommes amenés à prendre en compte des contraintes spatiales.

Le chapitre 3 porte sur la description des deux **approches actuellement utilisées** (la méthode de classement et le modèle Tau) pour simuler des structures géométriques sous contraintes spatiales. Leurs points forts et leurs points faibles sont étudiés, tant sur des modèles théoriques (e.g. modèle autodual, modèle gaussien trivariable) que sur des exemples simulés (agrégats d'objets, fonction gaussienne seuillée).

Les inconvénients des approches en cour nous amènent au chapitre 4 qui introduit l'**approche que nous avons développée**. Nous en donnons les principes généraux, en détaillons le fonctionnement et la comparons aux approches du chapitre 3. Ce chapitre est illustré par plusieurs exemples de validation qui mettent en jeu diverses hétérogénéités spatiales (proportion locale, orientation régionale ...)

Disposant d'une approche permettant de reproduire des structures géométriques complexes tout en intégrant des contraintes non stationnaires, nous arrivons maintenant au coeur de notre travail : simulation d'un modèle géologique (génétique ou non) sous contraintes spatiales. Le chapitre 5 présente un exemple d'**application de notre approche à un modèle génétique**.

Ce manuscrit se termine par une **conclusion** générale, qui résume les études effectuées, rappelle les résultats obtenus et donne quelques perspectives.

Enfin, figurent en **annexe** l'ensemble des notations, des concepts probabilistes et géométriques que nous avons utilisés dans cette thèse.

## 1.1 Réservoir pétrolier

Un réservoir pétrolier est un ensemble de dépôts sédimentaires interconnectés contenant des hydrocarbures et de l'eau. Ces dépôts sont constitués de roches poreuses et perméables, à l'intérieur desquelles les fluides peuvent circuler. Une barrière imperméable piégeant les fluides complète le réservoir. Les réservoirs pétroliers sont en général des milieux très hétérogènes, qui peuvent être fracturés et chenalisés.

Un certain nombre de mécanismes naturels facilitent la production, tels que l'expansion des fluides ou bien l'invasion du gisement par l'eau de la formation aquifère. Il y a quelques dizaines d'années, les compagnies pétrolières pouvaient encore espérer découvrir de nouveaux champs géants d'exploitation facile et rentable. Actuellement, la plupart des réserves mondiales de pétrole proviennent d'une meilleure exploitation de champs déjà connus par des méthodes de plus en plus sophistiquées telles que l'injection d'eau, de gaz ou de produits chimiques. Les réserves récupérables sont donc estimées en fonction des différents modes d'exploitation des champs.

Compte tenu de cette évolution, chaque compagnie pétrolière se doit de connaître de manière détaillée l'architecture interne de ses réservoirs souterrains. Bien souvent, les premières décisions à prendre lors du développement d'un projet pétrolier sont les plus difficiles. Il s'agit d'estimer la potentialité de production du champ à l'aide des quelques données disponibles. En raison des coûts de production élevés, l'industrie pétrolière a besoin de décrire la structure géologique des réservoirs et les propriétés des roches qui les constituent.

## 1.2 Modélisation numérique d'un réservoir

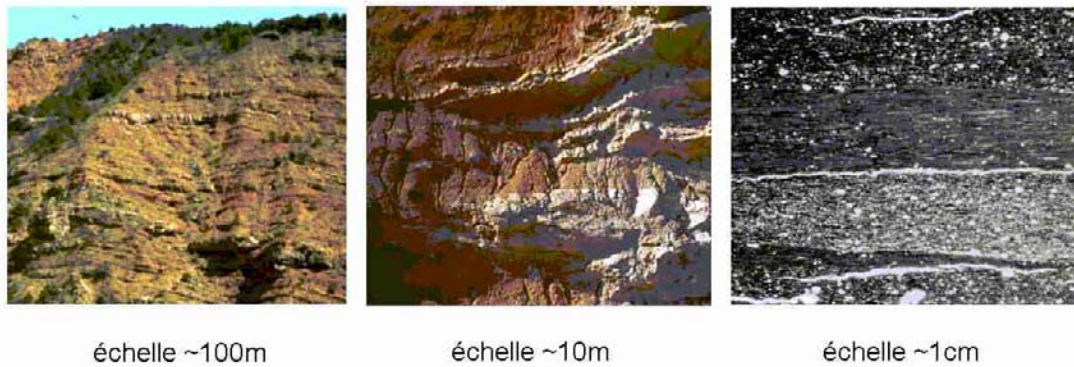
La caractérisation d'un réservoir est indispensable tout au long de son évolution, de sa découverte à son épuisement. C'est là où les modèles numériques représentent au fur et à mesure le gisement exploité. La construction d'un modèle numérique nécessite un minimum de connaissances sur le réservoir, notamment, les caractéristiques physiques, qui sont obtenues essentiellement à partir de forages de reconnaissance, de carottages et de diagraphies.

Naturellement, les caractéristiques physiques d'un réservoir varient continûment dans

l'espace et dans le temps. Mais les moyens numériques de modélisation sont limités par la capacité de calcul. C'est la raison pour laquelle le gisement est discrétisé en blocs, auxquels on attribue différents lithofaciès et des propriétés pétrophysiques numérisées. Expérimentalement, la modélisation d'un réservoir consiste à représenter numériquement ses hétérogénéités à différentes échelles. Une telle modélisation se fait selon les trois grandes étapes suivantes :

La **modélisation géologique** est la première étape de modélisation du réservoir. Elle consiste en une description numérique de la structure lithologique interne, c'est-à-dire des volumes constituant les différents objets géologiques (failles, fractures, lentilles, chenaux ...). Les hétérogénéités du réservoir dépendent des processus sédimentaires mettant en jeu le temps, la température et un certain nombre de conditions physico-chimiques. Elles peuvent donc être de nature et d'échelle différentes [Bour et al. (2002), Kazemi et al. (2005)] (Fig. 1).

Le modèle numérique géologique est souvent très détaillé et représente les objets géologiques contrôlant l'écoulement (la dimension de la maille de discrétisation est autour de 10m). La variabilité des structures géologiques est, en général, plus fine verticalement qu'horizontalement. Cette anisotropie spatiale impose l'anisotropie de discrétisation, i.e. la dimension horizontale de maille est souvent supérieure à sa dimension verticale (e.g. maille d'un modèle géologique en 3D peut être d'ordre de 10m, 10m et 1m en X, Y et Z).



**Fig. 1 Formations sédimentaires d'échelle différente.**

Cette première étape de modélisation est très importante, car la qualité de simulation du comportement d'un réservoir est fortement liée au réalisme du modèle géologique. Pour cette raison, il est indispensable de savoir intégrer des données de natures différentes (sismique, diagraphie, mesure de carotte ...) dès cette première étape de modélisation [Beucher (1998)].

L'étape suivante est la **modélisation pétrophysique**. Une fois les objets géologiques et leur répartition spatiale modélisés, il importe de spécifier les propriétés des roches qui les constituent [Bour et al. (2001), de Dreuzy et al. (2004)]. Les principales variables d'intérêt

sont la porosité, la perméabilité, la résistance de la roche. Ces propriétés sont définies sur le maillage géologique afin de mieux décrire leur variabilité.

La dernière étape de modélisation de réservoir est une **modélisation hydrodynamique**. Un modèle discrétisé du réservoir étant bâti, des simulations d'écoulement peuvent être envisagées [Sarda et al. (2001)]. Par souci d'économie de temps calcul, les modèles hydrodynamiques sont réalisés sur des grilles plus larges que les modèles géologiques (la maille d'un modèle hydrodynamique en 3D est typiquement d'ordre de 100m, 100m et 10m en X, Y et Z). Bien que jugé irréaliste par les géologues, ce modèle hydrodynamique s'avère néanmoins efficace pour simuler le comportement d'un réservoir et en prédire la production.

La qualité de la prédiction fournie par le modèle hydrodynamique dépend de nombreux facteurs. Ici nous nous bornons à n'en mentionner que les deux plus importants :

Le premier facteur à prendre en compte pour construire un modèle réaliste de réservoir est la reproduction de sa structure interne. Par exemple, la substitution d'un réseau de fractures par un milieu poreux dans le modèle géologique engendre une sous-estimation de la vitesse de fluide dans le modèle hydrodynamique. De son côté, la sous-estimation de vitesse de fluide provoque une sous-estimation du temps d'arrivée de l'eau au puits de production. La simulation des écoulements dans un tel contexte nécessite donc une très bonne représentation des hétérogénéités majeures que constituent les objets géologiques.

Le deuxième facteur, qui affecte largement la qualité de prédictions en production, est l'intégration des données dynamiques (calage de données d'essais de puits ou de production) qui est effectuée à cette étape de modélisation. Si même après avoir effectué un calage historique, les données obtenues par la modélisation hydrodynamique restent éloignées des données réelles, nous serons amenés à **modifier le modèle géologique** [Roggero et al. (2005)] afin de réduire cette différence et mieux représenter l'hétérogénéité souterraine.

## 1.3 Modélisation géologique

La construction d'un modèle géologique de réservoir est basée sur la représentation de l'environnement de dépôt géologique, y compris les hétérogénéités majeures contrôlant l'écoulement des fluides (chenaux fluviaux, couches imperméables, failles ...). La fiabilité de ce modèle augmente au fur et à mesure qu'il est conditionné aux données quantitatives telles que mesures de carottes, diagraphie, sismique.

On peut distinguer deux types d'approches de modélisation géologique : *l'approche génétique* et *l'approche stochastique*.

### 1.3.1 Approche génétique

Les premiers travaux entrepris pour modéliser la structure géologique interne d'un réservoir relèvent de cette approche. Ils s'appuient sur une modélisation des processus de sédimentation, permettant ainsi de respecter les hétérogénéités naturelles du réservoir. Cette modélisation est avant tout déterministe par prise en compte des lois physiques qui régissent l'évolution des milieux fluviaux à échelle géologique, mais présente aussi un aspect aléatoire pour reproduire le comportement parfois erratique de la nature. Les modèles génétiques des premiers travaux de Matheron (1969) ont été repris et poursuivis dans plusieurs directions.

Le **modèle d'architecture fluviale** résulte de la simulation des blocs gréseux représentant des ceintures de chenaux. Ce modèle est basé sur une formulation mathématique de concepts et de connaissances géologiques sans pour autant intégrer des processus sédimentaires. Le **modèle hydraulique** de bassin repose sur l'utilisation de lois hydrauliques à l'échelle du bassin, ce qui lui permet d'appréhender la géométrie générale du réservoir. Lorsqu'il s'appuie sur les lois physiques de sédimentation, ce type de modèle intègre de nombreux paramètres contrôlant la formation des dépôts. Le **modèle fluviale dynamique** porte à la fois sur la géométrie des systèmes fluviaux et sur leur évolution dynamique. La tendance à incorporer le maximum de processus physiques liés aux formations sédimentaires (transport hydraulique, déviation et éboulement des chenaux, migration latérale etc.) rend la modélisation génétique de plus en plus complexe. Pour plus de détail, le lecteur pourra consulter les travaux de Kolterman et Gorelick (1996) qui dressent un inventaire des différents types d'approche génétique.

Au cours des dernières années, des progrès considérables ont été réalisés dans la compréhension des systèmes sédimentaires méandriformes. Teles et al. (1998) et plus tard



Lopez (2003) proposent une modélisation génétique qui incorpore des processus de sédimentation ainsi que l'évolution tectonique de chenaux en temps géologiques pour restituer le système méandriforme dans toute sa complexité.

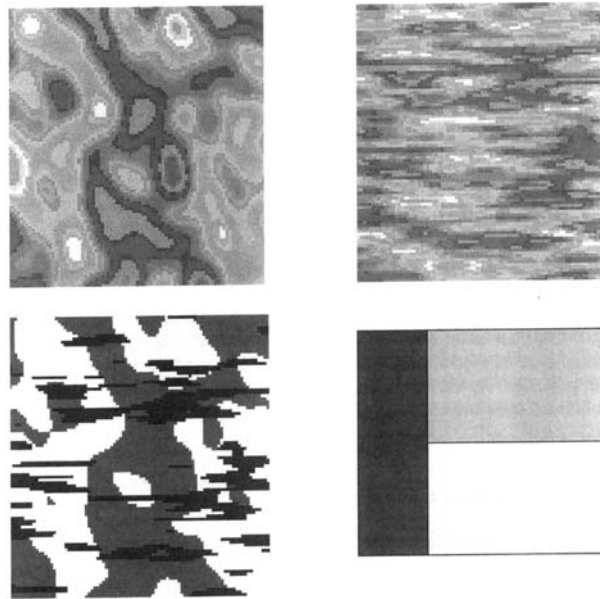
L'avantage majeur des modèles génétiques est qu'ils fournissent des résultats facilement interprétables en termes de processus sédimentaires. Ils conduisent également à des architectures géologiques beaucoup plus réalistes que les modèles stochastiques. Leur principal inconvénient est qu'il est difficile de les conditionner de manière satisfaisante à des données quantitatives telles que les données de puits ou de sismique. Le conditionnement d'un modèle contenant des objets géologiques paraît beaucoup plus facile si sa structure interne est définie par une fonction aléatoire. C'est ainsi que peu à peu apparaît l'approche stochastique.

### 1.3.2 *Approche stochastique*

L'approche stochastique consiste à concevoir une fonction aléatoire pour représenter l'hétérogénéité spatiale d'un réservoir. De nombreux modèles probabilistes peuvent être envisagés pour caractériser un réservoir [Matheron (1968), Chilès et Delfiner (1999), Lantuejoul (2002)]. Un modèle est choisi en fonction de l'environnement de dépôt, de l'échelle de travail et de la géométrie spécifique (connue à priori par les géologues). Parmi les modèles stochastiques d'utilisation courante, il est d'usage de distinguer deux grandes catégories : les modèles de type pixel et les modèles de type objet.

Les **modèles de type pixel** fournissent des réalisations d'une fonction aléatoire aux nœuds d'une grille, régulière ou non. En utilisant peu de paramètres (par rapport au modèle génétique et au modèle de type objet), le modèle de type pixel est capable de décrire de fortes hétérogénéités tout en intégrant des relations entre faciès géologiques (Fig. 2).

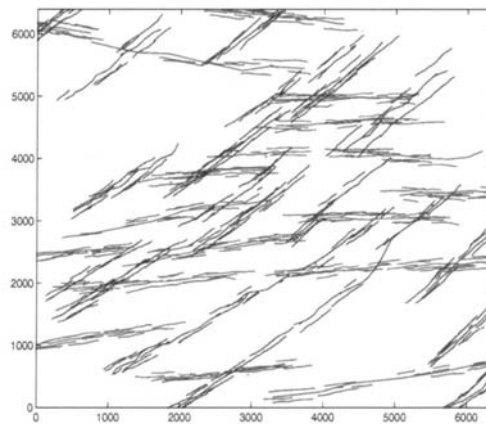
Il existe un vaste choix de modèles de type pixel. Son choix effectif est dicté par la structure géométrique à reproduire. Citons ici deux modèles d'utilisation courante : la fonction gaussienne seuillée [Matheron et al. (1987)] et la simulation séquentielle d'indicatrice [Journel et Alabert (1990)].



**Fig. 2 Exemple de la simulation plurigaussienne. Les images en haut sont gaussiennes. En bas à droite l'image représente la proportion et le schéma de troncature de 3 types de faciès. Le résultat de la troncature est en bas à gauche. [Lantuéjoul 2002].**

Les modèles de type pixel se prêtent bien au conditionnement par des données quantitatives, y compris des données dynamiques de production [Hu (2000)]. Souvent, des statistiques spatiales sont disponibles (moyenne, variance, variogramme ...) [Hu et Le Ravalec-Dupin (2004)]. Ces statistiques peuvent être utilisées pour ajuster la fonction aléatoire représentant le modèle afin d'améliorer la reproduction de la géométrie d'un milieu hétérogène. Mais, ni la diversité des modèles, ni leur ajustement ne permettent de capter et de reproduire certaines structures géologiques complexes telles qu'un système de chenaux méandriformes par exemple.

Sous l'appellation **modèles de type objet**, on désigne tout modèle qui utilise des formes géométriques plus ou moins complexes pour représenter les hétérogénéités d'un réservoir (chenaux [Deutsch et Wang (1996)], failles et fractures [Cacas et al. (2001), Gauthier et al. (2002)], crevasses, barrières etc.). Ces formes géométriques ou objets sont des représentations idéalisées des corps géologiques du réservoir. Chaque objet est défini par sa position et par ses caractéristiques géométriques (forme, taille, orientation). Des règles de répulsion ou d'attraction entre les corps simulés doivent être employées afin de décrire leur répartition dans l'espace. Un exemple (Fig. 3) d'utilisation de l'approche objet en milieu fracturé peut être vu chez Jenni (2005).



**Fig. 3 Exemple de modèle de type objet en 2D pour simuler un milieu fracturé [Jenni (2005)].**

En toute rigueur, par rapport aux modèles de type pixel, les modèles de type objet sont plus aptes à représenter des dépôts sableux d'origine fluviale, des réseaux de fractures, des milieux poreux hétérogènes. Cela confirme que par rapport aux modèles de type pixels, les modèles de type objet ont un caractère génétique plus affirmé. Moyennant quoi, le conditionnement de modèles de type objet aux données quantitatives n'est pas toujours facile. Par ailleurs, les données sismiques peuvent parfois témoigner d'une dérive le long du réservoir. Dans ce cas, les objets géologiques doivent être pris de loi non stationnaire et simulés en conséquence [Garcia-Morales (2003), Beucher et al. (2004)]. Le calage historique de tels modèles n'est pas toujours évident [Hu (2003), Jenni et al. (2004), Le Ravalec-Dupin et Hu (2004)].

Certaines modélisations stochastiques reposent sur une fonction aléatoire qui est explicitement spécifiée par sa loi spatiale. Cette loi spatiale doit être ajustée au moyen de quelques paramètres statistiques (moyenne, écart-type, variogramme). La géométrie produite par les réalisations de cette fonction aléatoire reste souvent éloignée de la géométrie réelle, spécialement dans le cas de réservoir de géométrie complexe fortement non stationnaire. D'autres modélisations stochastiques parent des données acquises aux puits ou bien de la sismique. Des paramètres expérimentaux en sont extraits afin de reproduire une géométrie plus réaliste. De façon générale, quelle que soit l'approche adoptée, plus la géométrie du réservoir est complexe, plus il est difficile de la conditionner aux données quantitatives.

De plus amples synthèses sur les différentes approches géostatistiques figurent par exemple dans les travaux de Dubrule (1998) et de Hu (2002).

### 1.3.3 Approche statistique multipoints

Une autre approche stochastique, récemment appliquée à la modélisation de réservoirs pétroliers, repose sur une statistique de reconnaissance de forme. Il s'agit de la reconstruction de structures géométriques complexes de façon aléatoire à partir d'une image d'apprentissage. Cette image est utilisée pour l'acquérir des lois probabilistes empiriques, lesquelles seront ensuite utilisées pour une simulation séquentielle.

L'idée d'utiliser une image d'apprentissage comme source de loi spatiale remonte à Guardiano et Strivastava (1993), puis a été développée par Caers et Journel (1998). Cette idée nous permet de nous affranchir du choix d'un modèle stochastique. **Une fonction aléatoire n'a plus besoin d'être définie explicitement.**

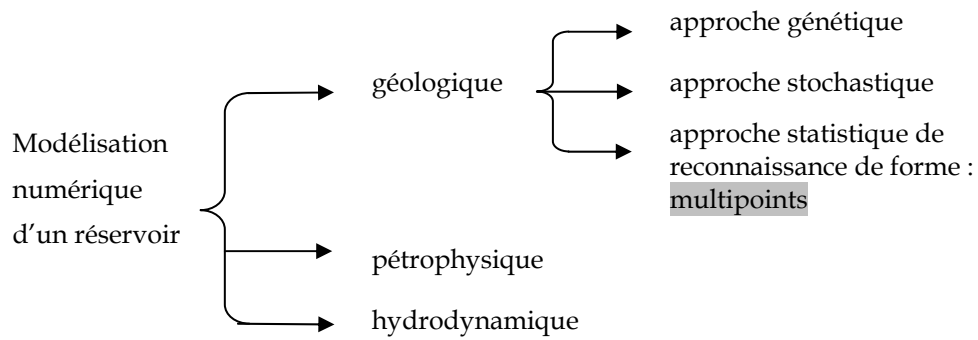
A l'heure actuelle, il existe deux algorithmes principaux de simulation liés à cette approche: par pattern [Arpat et Caers (2004b), Arpat et Caers (2004a), Zhang et al. (2004), Zhang et al. (2006)] et par pixel [Caers et al. (1999), Strebel et Journel (2001), Journel (2004)]. La méthode de simulation séquentielle par pixel (aussi connue comme l'approche **statistique multipoints**) se prête mieux au conditionnement [Akamine-Ramirez et Caers (2006)] que la simulation par pattern.

Pendant longtemps, l'approche multipoints (MP) a présenté l'inconvénient d'être particulièrement gourmande en mémoire et en temps d'exécution. Pour remédier à ces problèmes, Strebel (2000) a développé un algorithme séquentiel, qui permet de stocker toutes les données statistiques inférées de l'image d'apprentissage dans une structure arborescente. Son algorithme a rendu l'approche multipoints opératoire [Liu et al. (2002), Maharaja (2004a), Feyen et Caers (2004), Okabe et Blunt (2004)].

Nous pouvons citer les avantages majeurs de l'approche MP :

- 1) souplesse de conditionnement [Strebel (2002)] ;
- 2) reproduction de formes géométriques complexes ;
- 3) consommation modérée en mémoire et en temps calcul.

Cette méthode sera présentée plus en détail au chapitre 2. Voici le schéma qui indique la place de l'approche *multipoints* dans la hiérarchie des méthodes de modélisation numérique d'un réservoir :



L'approche MP étant basée sur une image d'apprentissage, le résultat de simulation dépend intégralement du contenu de cette image. En pratique, cette image d'apprentissage est simulée selon un modèle stochastique (de type objet ou de type pixel), ce qui ne permet pas toujours de tenir compte de la nature sédimentaire du réservoir pétrolier [Strebel et al. (2002), Harding et al. (2004), Liu et al. (2004)]. Au paragraphe suivant, l'idée d'une modélisation génétique MP est introduite.

## 1.4 Cadre de la thèse

Prises séparément, chacune des grandes approches (génétique et stochastique) ne parvient pas à répondre de façon satisfaisante aux exigences de la modélisation d'un réservoir pétrolier : 1) reproduire la structure géologique sédimentaire et ses hétérogénéités de façon la plus réaliste possible ; 2) intégrer l'ensemble des données d'origines diverses et d'échelles différentes ; 3) évaluer les incertitudes par génération de multiples scénarios équiprobables pour leur utilisation subséquente. L'objectif de cette thèse est de proposer une solution à la fois adéquate et opératoire (du point de vue temps calcul) à ces exigences.

Il se pourrait que la solution se réside dans une **approche hybride**.

La première étape de modélisation géologique d'un réservoir est envisagée par approche génétique. Cette approche permet de reproduire l'environnement de dépôt géologique et représenter les hétérogénéités majeures qui contrôlent l'écoulement des fluides ; cependant le modèle génétique reste non contraint aux données quantitatives. Pour cette raison, une des réalisations du modèle génétique prendra le rôle d'image d'apprentissage. La deuxième étape de modélisation sera alors effectuée par approche statistique de reconnaissance de forme appliquée à cette image d'apprentissage. On peut aussi espérer de reproduire des formes géométriques complexes tout en intégrant des données et des contraintes disponibles. Puisque la simulation est stochastique, des multiples réalisations peuvent être obtenues.

Dans cette thèse, nous supposons que nous disposons d'une ou de plusieurs réalisations du modèle génétique d'un réservoir (images d'apprentissage). Nous supposons également, que la structure interne du réservoir est de géométrie complexe et non stationnaire. Notre objectif est **la simulation d'une structure géométrique complexe en intégrant la non stationnarité spatiale**. Selon l'intérêt potentiel de l'approche MP [Caers et al. (1999), Caers et Zhang (2002), Caumon et al. (2004); Caumon et Journel (2006)], nous déciderons de l'exploiter afin de réaliser notre objectif.

Le point de départ de cette thèse est d'étudier l'algorithme de simulation MP, avec l'idée de voir comment prendre en compte des contraintes non stationnaires. Après avoir étudié l'approche MP dans un cadre stationnaire, nous mènerons une étude critique des approches de simulation MP non stationnaire publiées dans la littérature, puis nous chercherons à élaborer une approche alternative permettant d'éviter les inconvénients des algorithmes existants.

L'approche développée va consister à incorporer des contraintes, venant généralement de

la sismique, et révélant une non stationnarité de la structure interne du réservoir, telles que la proportion locale des objets géologiques, leur taille, leur orientation régionale etc.

## 2 Approche statistique multipoints

Dans ce chapitre nous présentons l'approche statistique multipoints et discutons ses avantages et ses inconvénients. Nous commençons par en donner le principe ainsi que l'algorithme initial proposés par Guardiano et Strivastava (1993). Nous décrivons ensuite la mise en œuvre de cette approche en détaillant la méthode de stockage des statistiques empiriques et de leur réutilisation proposée par Strebelle (2002). Une attention particulière sera portée aux conditions limites d'application de l'approche MP. Soulignons que tous les exemples présentés dans ce chapitre sont effectués par simulation stationnaire. En fin de chapitre, nous abordons le problème central de cette étude – la simulation non stationnaire.

### 2.1 Présentation générale

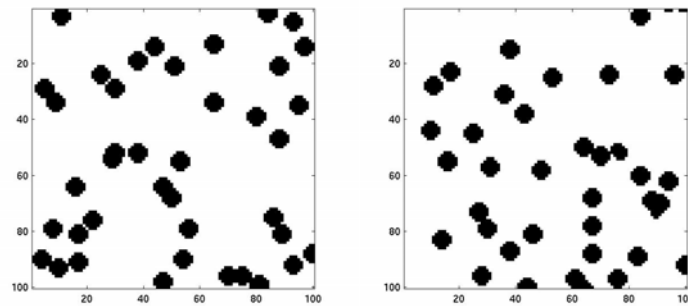
#### 2.1.1 Principe

Le principe de base de l'approche MP est d'utiliser l'image d'apprentissage, supposée représentative de la structure géométrique de réservoir, comme la source des statistiques empiriques d'ordre multiple. À partir de ces statistiques, la loi spatiale d'une fonction aléatoire est construite. Cette loi permet de reproduire une structure géométrique complexe. L'algorithme de simulation MP s'appuie ainsi sur trois idées :

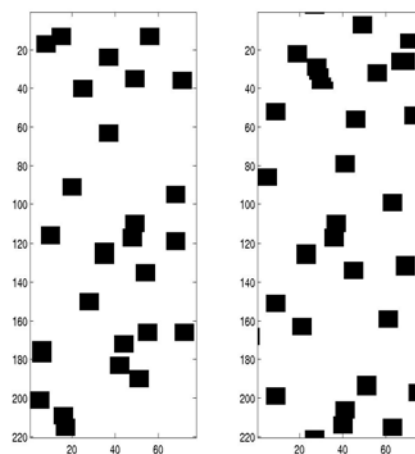
- 1) utilisation d'une image d'apprentissage ;
- 2) acquisition des statistiques MP à partir de cette image;
- 3) simulation séquentielle à partir des statistiques MP acquises.

L'algorithme ne mettant pas en jeu de forme géométrique de façon explicite, leurs caractéristiques (taille, orientation, longueur etc.) doivent être reproduites de façon implicite. Les exemples présentés (Fig. 4-7) ont été choisis de natures différentes (modèle booléen, modèle gaussien, réseau de fractures) pour illustrer les possibilités de l'approche multipoints.

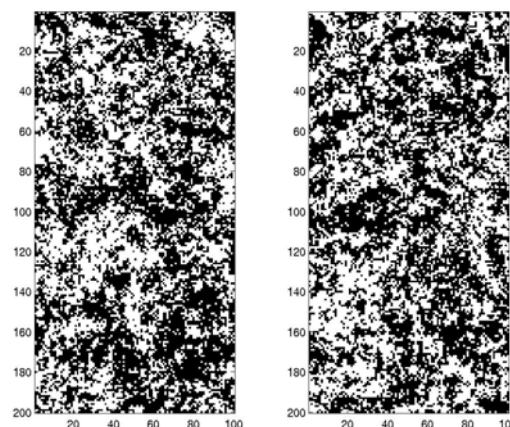




**Fig. 4** Simulation MP stationnaire. Reproduction d'un modèle booléen avec des objets ronds : l'image d'apprentissage à gauche, la réalisation MP à droite.



**Fig. 5** Simulation MP stationnaire. Reproduction d'un modèle booléen avec des objets carrés : l'image d'apprentissage à gauche, la réalisation MP à droite.



**Fig. 6** Simulation MP stationnaire. Reproduction d'un modèle gaussien seuillé (variogramme hyperbolique avec une portée unité) : l'image d'apprentissage à gauche, la réalisation MP à droite.

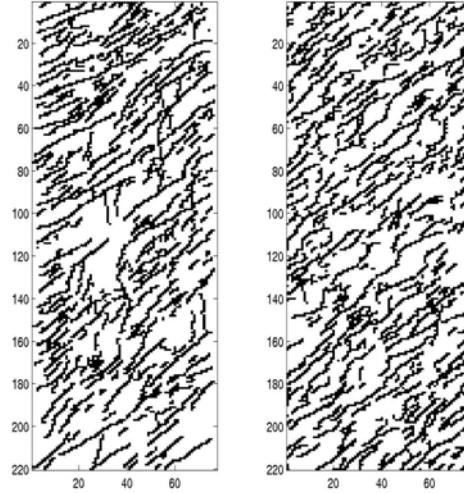


Fig. 7 Simulation MP stationnaire. Reproduction d'un réseau de fractures en 2D : l'image d'apprentissage à gauche, la réalisation MP à droite.

### 2.1.2 Algorithme initial

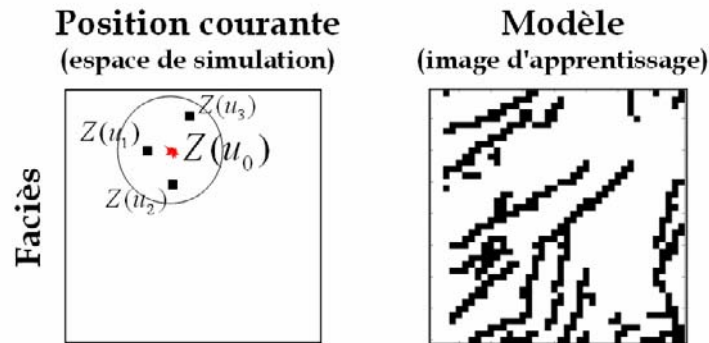


Fig. 8 L'espace de simulation à gauche ; l'image d'apprentissage à droite.

Ramenée à un cadre discret, l'approche MP se présente de la manière suivante :

Soit  $\mathbf{I}$  la grille de discrétisation de l'Image d'apprentissage, et soit  $\mathbf{E}$  la grille de l'Espace de simulation. Usuellement,  $\mathbf{I}$  et  $\mathbf{E}$  ont la même dimension physique (2 ou 3) et ont la même résolution spatiale. Par contre, leur taille peut être très différente. À chaque point  $u$  de l'une de deux grilles est attaché, ou bien doit être attaché, un attribut discret  $Z(u)$  (fonction aléatoire) qui peut prendre  $K$  états correspondants aux  $K$  faciès différents. La simulation MP consiste à générer une valeur d'attribut en chaque point de  $\mathbf{E}$ .

Supposons que l'on veuille simuler  $Z$  au point  $u_0 \in E$  (Fig. 8). Soient  $u_1, u_2, \dots, u_n \in E$  un ensemble des points ordonnés composé par les points déjà simulés au tour de  $u_0$  (Fig. 9).

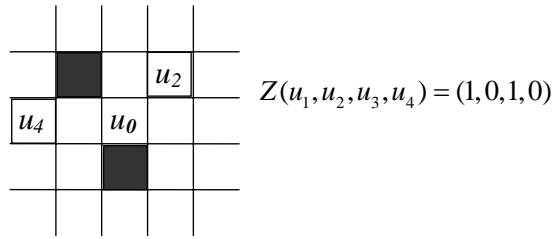
Déf. On appelle voisinage de  $u_0$  toute famille finie de points ordonnés  $\{u_1, u_2, \dots, u_n\}$  ne contenant pas  $u_0$ . Cet ensemble de points sera noté  $\mathbf{u}$  dans ce qui suit.

Déf. Le voisinage  $\mathbf{u}$  du point  $u_0$  est équivalent au voisinage  $\mathbf{w}$  du point  $w_0$  s'ils ont le même nombre de points et s'ils coïncident à une translation près :

$$\mathbf{u} \equiv \mathbf{w} \Leftrightarrow (u_1 - u_0, u_2 - u_0, \dots, u_n - u_0) = (w_1 - w_0, w_2 - w_0, \dots, w_n - w_0).$$

Soit  $\mathbf{z} = (z_1, z_2, \dots, z_n)$  un *état* (une réalisation) de  $Z$  sur un voisinage  $\mathbf{u}$ . On écrit  $Z(u_1, u_2, \dots, u_n) = (z_1, z_2, \dots, z_n)$  ou, simplement,  $Z(\mathbf{u}) = \mathbf{z}$ .

Déf. Dire que  $Z(\mathbf{w}) = \mathbf{y}$  est une *réplique* de  $Z(\mathbf{u}) = \mathbf{z}$  est signifie l'identité des voisinages  $\mathbf{w} \equiv \mathbf{u}$  et l'égalité des valeurs  $\mathbf{y} = \mathbf{z}$ .



**Fig. 9 Un voisinage simulé (*état*) du point  $u_0$ . Cas 2D.**

Les valeurs de  $Z(\mathbf{I})$  étant structurées, la valeur d'un point  $w_0 \in \mathbf{I}$  dépend des valeurs de ses points voisins  $\mathbf{w} \subset \mathbf{I}$ . Pour pouvoir reproduire la structure, la valeur en un point  $u_0 \in E$  est générée en fonction des points déjà simulés de son voisinage  $\mathbf{u} \subset E$ . Ainsi, pour simuler la variable aléatoire  $Z(u_0)$ , il nous faut estimer la loi conditionnelle  $P[Z(u_0) | Z(\mathbf{u}) = \mathbf{z}]$ . Cette loi est extraite de l'image d'apprentissage en enregistrant l'ensemble des répliques de  $Z(\mathbf{u}) = \mathbf{z}$ . La valeur de  $u_0$  est ensuite générée selon cette loi conditionnelle discrète (Fig. 10). La simulation s'arrête quand tous les points de l'espace de simulation ont été traités. L'algorithme de simulation MP s'écrit alors ainsi :

- (i) Choisir un point non encore simulé  $u_0 \in E$  et en extraire l'état de son voisinage  $Z(\mathbf{u}) = \mathbf{z}$  ;
- (ii) Enregistrer toutes les répliques de  $Z(\mathbf{u}) = \mathbf{z}$  dans l'image d'apprentissage :  
 $W = \{w_0 \in \mathbf{I} : Z(\mathbf{w}) = \mathbf{z}, \mathbf{w} \equiv \mathbf{u}\}$  ;
- (iii) Si  $W = \emptyset$ , faire  $\mathbf{u} = \mathbf{u} \setminus u_n$  et retourner en (ii) ;

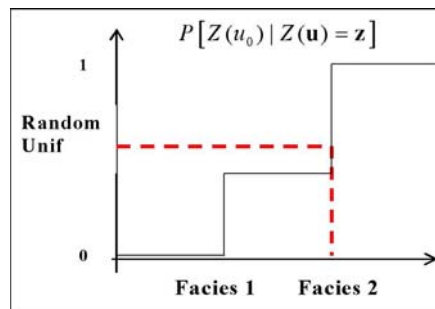
(iv) Simuler la valeur de  $Z$  au point  $u_0$  selon la loi conditionnelle :

$$P[Z(u_0) = k \mid Z(\mathbf{u}) = \mathbf{z}] = \frac{P[Z(u_0) = k, Z(\mathbf{u}) = \mathbf{z}]}{P[Z(\mathbf{u}) = \mathbf{z}]} .$$

$$= \frac{\#\{w_0 \in W : Z(w_0) = k\}}{\#\{w_0 \in W\}}, \quad k = 1, 2, \dots, K$$

où le symbole  $\#$  désigne la cardinalité ensembliste.

(v) Retourner en (i).



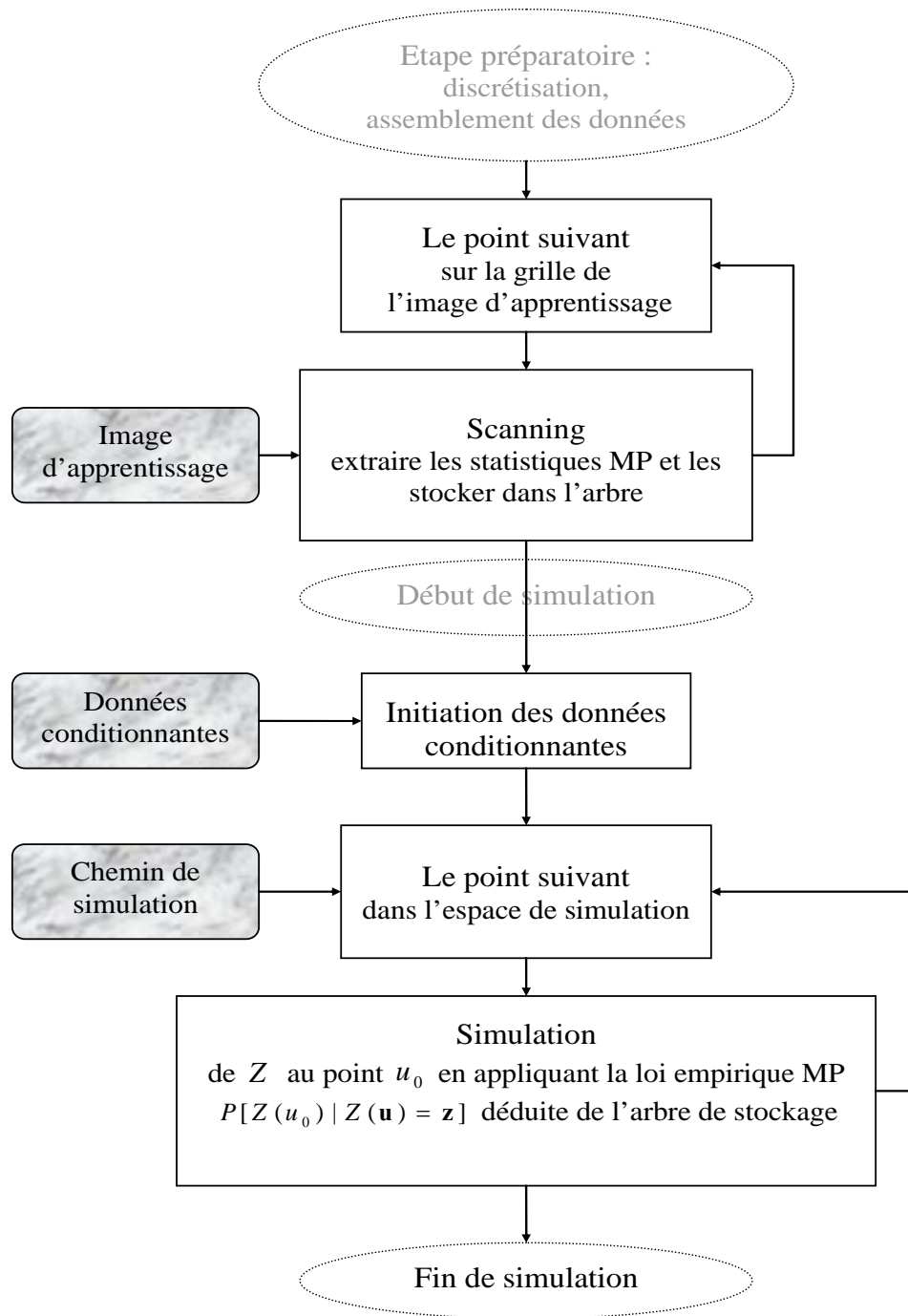
**Fig. 10** Simulation d'une fonction aléatoire à partir d'une loi empirique pour deux faciès.

## 2.2 Mise en oeuvre

L'algorithme, initialement proposé par Guardiano et Strivastava (1993), présente l'inconvénient d'être particulièrement gourmand en mémoire et en temps calcul. En effet, le scanning de l'image d'apprentissage est nécessaire pour la simulation de chaque point. Strebelle (2000) a proposé un algorithme permettant de ne scanner qu'une seule fois l'image d'apprentissage, et de mémoriser les statistiques multipoints sous une forme arborescente (nommée ensuite *arbre de stockage*) afin de pouvoir en disposer plus rapidement.

Une fois l'image d'apprentissage scannée, la simulation peut débuter. Les points sont simulés selon un chemin prédéfini (aléatoire ou non) qui parcourt tout l'espace de simulation. Pour chaque point traité, une loi conditionnelle est extraite directement de l'arbre de stockage, correspondant à l'état de son voisinage. La simulation s'arrête quand tous les points du champ de simulation ont été traités. Néanmoins, il existe dans la littérature des algorithmes MP de post-simulation, qui consistent à resimuler des points afin de raffiner la reproduction de la géométrie. Nous discuterons de cette problématique au paragraphe 2.3.3. Dans le cadre de cette thèse, pour ne pas trop nous disperser, nous utilisons l'algorithme MP où chaque point n'est simulé qu'une seule fois.

Voici un diagramme de l'algorithme MP (Fig. 11).



**Fig. 11** Algorithme de simulation statistique MP.

### 2.2.1 Scanning de l'image d'apprentissage

L'étape de scanning de l'algorithme MP consiste à dériver et stocker les statistiques MP au moyen d'une fenêtre glissante de taille fixe (dite *template* dans la littérature anglophone), qui explore toute l'image d'apprentissage point après point (Fig. 12). À l'intérieur de cette fenêtre, chaque état sera enregistré dans l'arbre de stockage.

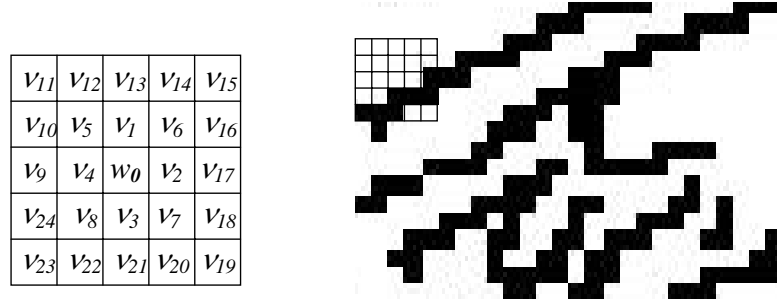
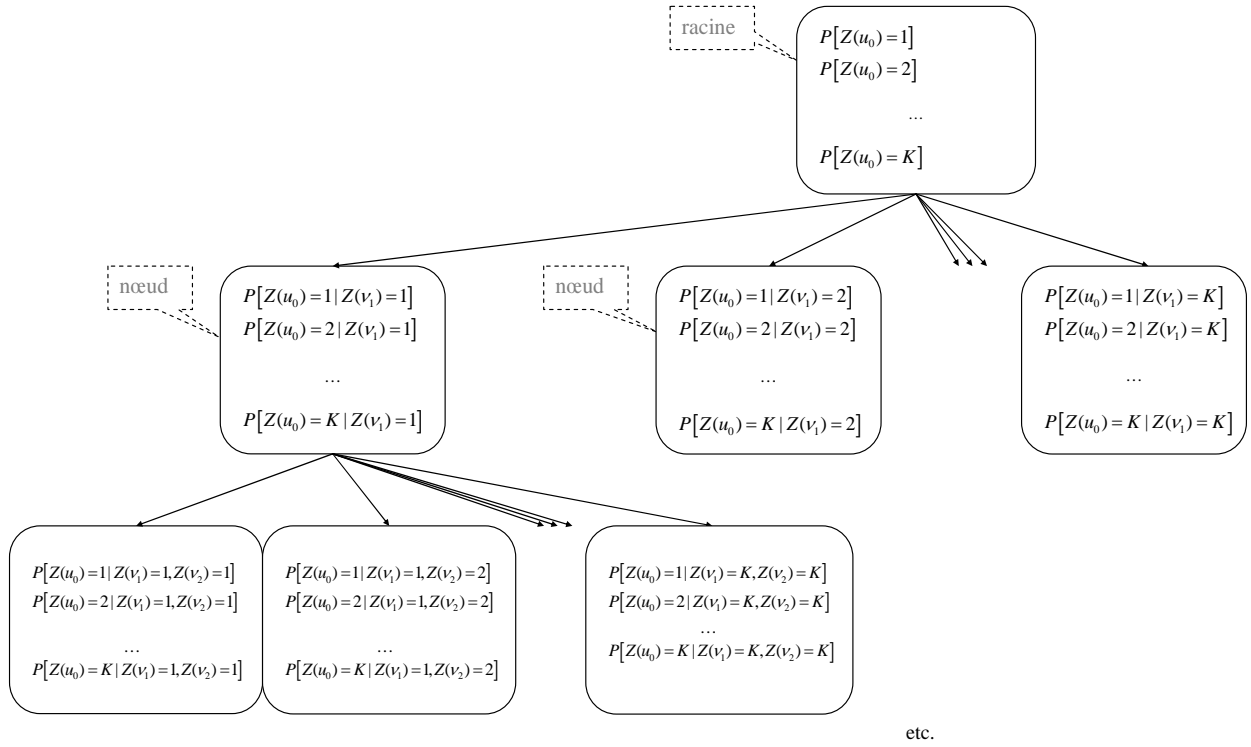


Fig. 12 Fenêtre d'exploration avec un point central  $w_0$ . Cas 2D.

Soit  $\mathbf{v}^l = (v_1, v_2, \dots, v_l)$ ;  $l: 1, \dots, L$ , un voisinage du point  $w_0$  qui contient l'ensemble de points les plus proches de  $w_0$ ,  $L$  étant le nombre maximal de ces points. La configuration de cet ensemble définit la zone de conditionnement du point  $w_0$ , c'est-à-dire que seules les valeurs de cet ensemble de points seront prises en compte pour dériver la loi conditionnelle de  $Z(w_0)$ . La numérotation de points dans un tel ensemble définit l'ordre de prise en compte des points voisins. Nous appelons un ensemble  $\mathbf{v}^L$  une *fenêtre*, des sous-ensembles  $\mathbf{v}^l, l < L$  — des *sous-fenêtres* et  $w_0$  — le *point central* de fenêtre. Les valeurs prises par la fonction aléatoire sur une fenêtre se dénotent par  $Z(v_1, v_2, \dots, v_l) = (z_1, z_2, \dots, z_l)$  ou, simplement,  $Z(\mathbf{v}^l) = \mathbf{z}^l$ .

Pendant le scanning le point central se déplace sur toute l'image d'apprentissage et enregistre les états de la fenêtre et de ses sous fenêtres dans l'arbre de stockage. Cet arbre est constitué des noeuds de niveau entre 0 et  $L$  (la taille maximale d'une fenêtre d'exploration a été préalablement choisie pour scanner l'image d'apprentissage). Un niveau  $l$  correspond à la sous fenêtre  $\mathbf{v}^l$  et peut contenir autant de noeuds qu'il existe d'états différents de  $Z(\mathbf{v}^l)$  dans l'image d'apprentissage. A chaque noeud de l'arbre de stockage sont associés 1) un état particulier  $Z(\mathbf{v}^l) = \mathbf{z}^l$ ; 2) les nombres de répliques de cet état pour les différentes valeurs du point central. L'ensemble de cette information est équivalent à une loi empirique  $P[Z(u_0) | Z(\mathbf{v}^l) = \mathbf{z}^l]$ ,  $u_0 \in I$ .



Le schéma de l'arbre de stockage est présenté dans le cas général de  $K$  faciès (Fig. 13). La Fig. 14 montre l'exemple d'un noeud avec quatre voisins définis dans le cas des deux faciès.



**Fig. 13** L'arbre de stockage pour les statistiques MP extraites de l'image d'apprentissage. Cas général de  $K$  faciès.

Nœud de niveau 4  $\{Z(v_1)=0, Z(v_2)=1, Z(v_3)=1, Z(v_4)=0\}$

Deux faciès :

$Z = 0$    
 $Z = 1$  

Nous enregistrons les statistiques suivantes :

$$P[Z(u_0)=0 | Z(v_1)=0, Z(v_2)=1, Z(v_3)=1, Z(v_4)=0]$$

$$P[Z(u_0)=1 | Z(v_1)=0, Z(v_2)=1, Z(v_3)=1, Z(v_4)=0]$$

	$v_5$	$v_l$	etc.	
	$v_4$	$u_0$	$v_2$	
		$v_3$		

**Fig. 14** Exemple d'un nœud de niveau  $l = 4$  pour deux faciès.

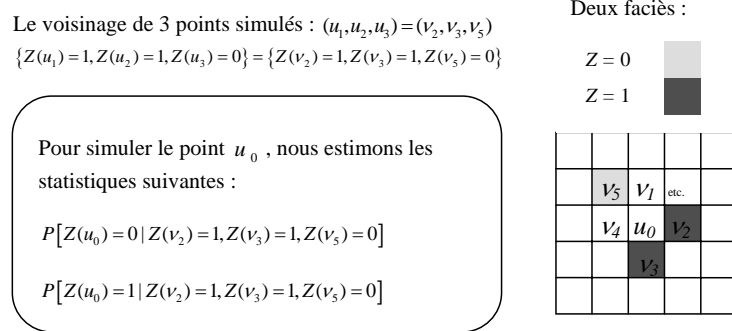
## 2.2.2 Simulation

La simulation commence par le positionnement des données fixes et certaines dans l'espace de simulation. (En pratique, ce sont les données de puits dont la position et les faciès géologiques sont connus.) Au cours de la simulation, chaque point simulé devient lui-



même un point fixe. Les points déjà évalués (d'une manière ou d'une autre) sont considérés comme des points simulés.

Soit  $\mathbf{u}$  le voisinage simulé du point  $u_0$ . Ce voisinage est un sous ensemble de fenêtre d'exploration, mais pas nécessairement une sous fenêtre  $\mathbf{u} \subseteq \mathbf{v}^l$ ,  $\mathbf{z} \subseteq \mathbf{z}^l$ . Ainsi, le nombre maximal des points simulés qui seront prises en compte pendant la simulation est limité par la taille de fenêtre d'exploration. L'exemple d'un voisinage simulé de trois points est présenté sur la Fig. 15.



**Fig. 15 Exemple d'un voisinage simulé pour deux faciès.**

En chaque point de simulation  $u_0 \in \mathbf{E}$ , il est nécessaire de connaître la loi conditionnelle :

$$P[Z(u_0) | Z(\mathbf{u}) = \mathbf{z}] \quad \text{Eq 2.1}$$

Pour estimer cette loi, on considère l'algorithme suivant :

- (i) Choisir le point suivant  $u_0 \in \mathbf{E}$  selon le chemin prédéfini et extraire l'état de son voisinage  $Z(\mathbf{u}) = \mathbf{z}$
- (ii) Choisir le niveau d'arborescence qui est la taille de la plus petite sous fenêtre contenant le voisinage simulé :

$$l = \min_{0 \leq l \leq L} \{\mathbf{u} \subseteq \mathbf{v}^l\}.$$

- (iii) Parmi l'ensemble des nœuds de ce niveau, retenir ceux qui coïncident avec l'état du voisinage  $Z(\mathbf{u}) = \mathbf{z}$  :

$$W = \{w_0 \in \mathbf{I} : Z(\mathbf{v}^l) = \mathbf{z}^l, \mathbf{z} \subseteq \mathbf{z}^l\}.$$

- (iv) Si  $W = \emptyset$ , faire  $\mathbf{u} = \mathbf{u} \setminus v_l$  et retourner en (ii) ;
- (v) Calculer la loi conditionnelle :

$$P[Z(u_0) = k | Z(\mathbf{u}) = \mathbf{z}] = \frac{\#\{w_0 \in W : Z(w_0) = k\}}{\#\{w_0 \in W\}}, \quad k = 1, 2, \dots, K \quad \text{Eq 2.2}$$

- (vi) Retourner en (i).

## 2.3 Conditions d'application

Les conditions d'application de l'approche MP sont dictées par des considérations théoriques, économique et pratiques.

### 2.3.1 Taille de l'image d'apprentissage

L'approche MP consiste à estimer la probabilité de  $Z(u_0)$  conditionnée à l'état du voisinage  $Z(\mathbf{u})$  par le nombre des répliques de cet état (Eq 2.2). D'après la loi forte des grands nombres, plus on retrouve ces répliques dans l'image d'apprentissage, plus notre estimation sera précise. Naturellement, la chance de trouver plusieurs répliques sur un voisinage de taille proche à celle de l'image d'apprentissage est très faible. Ainsi, nous supposons, que la taille de l'image d'apprentissage (nombre de points au total) est relativement grande par rapport à la taille de fenêtre d'exploration (nombre maximal de voisinage).

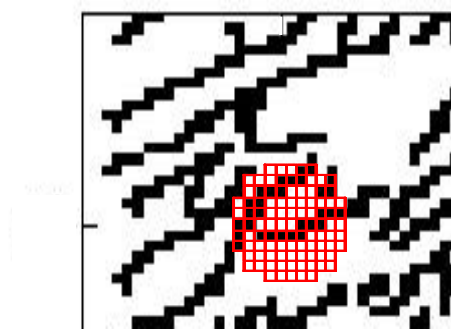
Étant donnée une image d'apprentissage de  $K$  faciès, nous voulons dériver les probabilités d'apparition d'une combinaison de  $n$  points voisins. Si l'image est complètement désordonnée (entropie forte), nous pouvons avoir  $K^n$  combinaisons différentes au maximum. Puisque nous espérons trouver plusieurs répliques de chaque combinaison dans l'image d'apprentissage, la taille de l'image doit être au moins supérieure au nombre de combinaisons possibles.

Prenons un simple exemple de structure 2D de deux faciès que nous voudrions explorer avec une fenêtre de  $10 \times 10 = 100$  points. Au pire nous aurons  $2^{100} \approx 10^{30}$  combinaisons différentes. Dans ce cas, il nous faut une image de la taille d'au moins de  $10^{15} \times 10^{15}$  points (pour avoir une réplique pour chaque combinaison possible), ce qui n'est jamais le cas en pratique.

Dans les faits, nous disposons de moins de  $10^{30}$  combinaisons différentes, car l'image d'apprentissage est structurée (entropie faible). Le travail de Arpat et al. (2002) montre un exemple de reconstruction d'un milieu fluvial avec seulement 100 patterns (combinaisons de 40 points chacune) les plus fréquents. De plus, il faut tenir compte que la simulation MP n'assure pas l'utilisation de toutes les statistiques MP dérivées de l'image d'apprentissage. En effet, une des réalisations MP ne présente pas tous les patterns se trouvant dans l'image d'apprentissage.

### 2.3.2 Fenêtre d'exploration

Le choix d'une fenêtre d'exploration adaptée est essentiel pour la reproduction correcte de la structure géométrique. La taille de fenêtre d'exploration doit être suffisamment grande par rapport à la taille caractéristique de la texture de l'image d'apprentissage. De plus, pour saisir puis reproduire l'étendue et l'agencement des faciès, la fenêtre d'exploration doit pouvoir couvrir à la fois un faciès qui représente l'objet géologique et un autre ou plusieurs autres pour représenter son environnement (cf. Fig. 16 dans le cas de deux faciès).



**Fig. 16 Exemple d'une fenêtre d'exploration qui couvre le faciès noir entouré par de faciès blanc.**

Dans le cadre de cette étude, les simulations MP ont été effectuées avec différentes tailles de fenêtre d'exploration. Il a été observé qu'au-delà d'une certaine taille du voisinage assez grande, le résultat de simulation n'est plus sensible à l'augmentation ultérieure de voisinage.

Il est possible d'adapter la forme d'une fenêtre d'exploration en fonction de l'orientation majeure des patterns. Pour étudier la sensibilité de la simulation MP vis-à-vis du choix de la forme de la fenêtre, Barrera et al. (2004) a effectué un test avec des fenêtres d'orientation différente. Ce test montre que les statistiques multipoints acquises par une fenêtre qui a la même orientation que les patterns sont plus robustes. Dans ce travail nous trouvons des indications sur le choix optimal de la forme d'une fenêtre. Le choix d'une fenêtre peut jouer un rôle important dans la reproduction de la structure si la taille de la fenêtre est optimale. Dès que la fenêtre dépasse une taille critique (dans toutes les dimensions de simulation), la forme n'est plus importante. Cela vient de ce que les statistiques MP acquises avec un voisinage assez grand sont capables de saisir toutes les caractéristiques géométriques de l'image, telles que la forme ou l'orientation.

Dans la suite de notre étude, nous supposons la fenêtre d'exploration de taille suffisamment grande. Sauf précision contraire, la fenêtre utilisée dans les simulations en

2D sera de taille 100 points, ordonnée comme suit (Fig. 17) :

				100	92	77	85	93				
	84	76	68	60	45	53	61	69	81			
	75	52	44	36	21	29	37	49	70			
99	67	43	28	20	9	13	25	38	62	94		
91	59	35	19	8	1	5	14	30	54	86		
80	48	24	12	4	0	2	10	22	46	78		
90	58	34	18	7	3	6	15	31	55	87		
98	66	42	27	17	11	16	26	39	63	95		
	74	51	41	33	23	32	40	50	71			
	83	73	65	57	47	56	64	72	82			
				97	89	79	88	96				

Fig. 17 Fenêtre d'exploration pour la simulation en 2D.

### 2.3.3 Chemin de simulation

Il y a plusieurs possibilités pour effectuer une simulation séquentielle.

Une des possibilités est de définir un **chemin aléatoire** qui parcourt tout l'espace de simulation indépendamment de la géométrie à simuler. Comme le nombre de points simulés augmente au fur et à mesure que la simulation se construit, les voisinages deviennent de plus en plus grands, si bien que les lois conditionnelles deviennent de plus en plus déterministes (Fig. 18). Ainsi, la probabilité conditionnelle à estimer dépend de toutes les estimations précédentes [Daly (2004)], et donc du chemin de simulation.

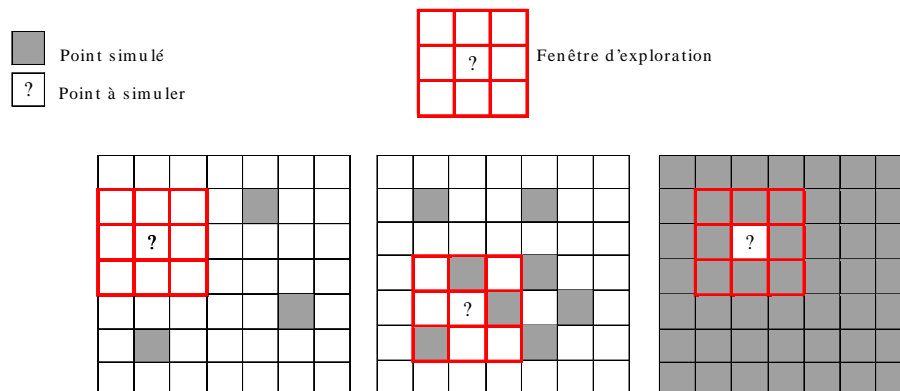


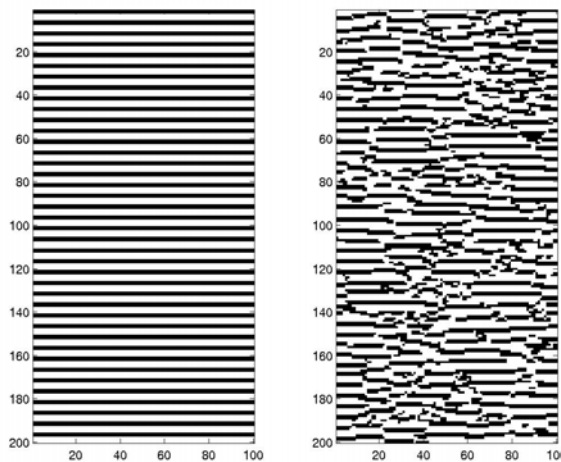
Fig. 18 Simulation par le chemin aléatoire (2D).

Soit  $u_1, u_2, \dots, u_{n-1}, u_n$  un chemin aléatoire de simulation. Le voisinage évolue à chaque nouveau point simulé. Ainsi, la probabilité conditionnelle du point  $u_n$  dépend des

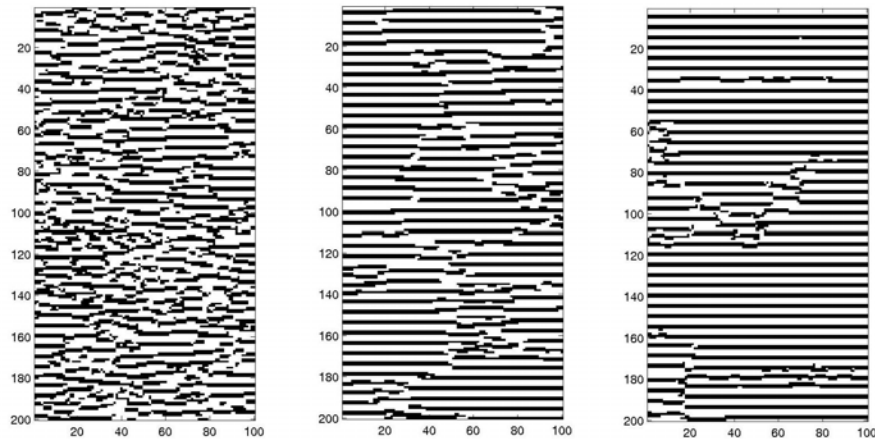
probabilités conditionnelles des points  $u_1, u_2, \dots, u_{n-1}$  de manière itérative.

Cet effet n'est pas négligeable dans le cas de la simulation d'objets continus (e.g. chenaux). Souvent, le manque de points conditionnants en début de simulation crée des configurations de points qui n'existent pas dans l'image d'apprentissage (le point blanc préalablement simulé sur le parcours d'un chenal), ce qui empêche la continuité des objets. Ici nous illustrons la non continuité causée par un cheminement aléatoire dans le cas de simulation MP de lignes parallèles (Fig. 19).

Pour une meilleure reproduction des objets de la taille supérieure à la taille de fenêtre d'exploration il est conseillé [Strebelle 2002] de simuler d'abord sur une grille grossière puis d'affiner la grille (technique **multigrille**). Ceci permet de reproduire la structure spatiale à une distance supérieure de celle de la fenêtre d'exploration (effet de nid, fractales). Nous proposons ici deux exemples de simulation qui montrent l'amélioration obtenue par des techniques multigrille d'ordre 3 et 5 (Fig. 20). Nous pouvons remarquer que la continuité des lignes restaurées est « proportionnelle » à l'ordre de multigrille.

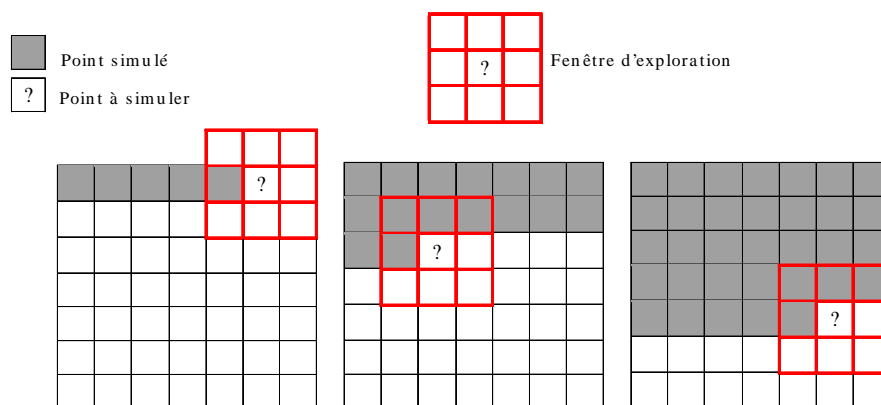


**Fig. 19 Simulation MP par le chemin aléatoire 2D. Reproduction des lignes parallèles : l'image d'apprentissage à gauche, la réalisation MP à droite.**

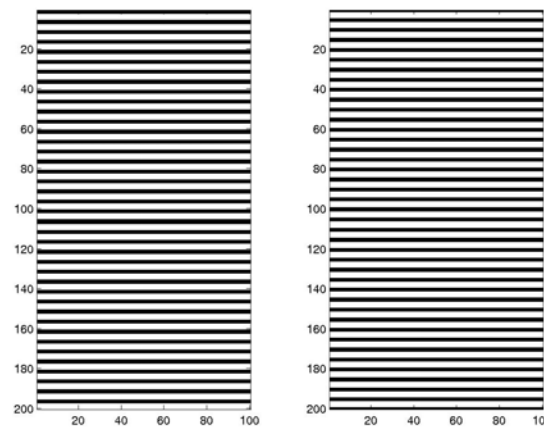


**Fig. 20 Simulation MP par le chemin aléatoire 2D : grille fine à gauche ; multigrille d'ordre 3 au centre ; multigrille d'ordre 5 à droite.**

Une autre façon de simuler séquentiellement est de définir un **chemin unilatéral** [Daly et Knudby (2007)]. Un exemple est un cheminement régulier selon l'ordre des points de la grille (Fig. 21). Dans ce cas, chaque point à simuler a son voisinage de points déjà simulé toujours de la même forme (sauf au bord). Il est alors simulé selon la même loi conditionnelle, au contraire des points d'un chemin aléatoire. Étant donnée une image d'apprentissage et une grille de simulation illimitées, nous nous trouvons dans le cadre du Modèle Markovien Unilatéral sans configurations inexistantes. En pratique toutefois, ce mode de simulation n'exclut pas des artefacts. Il s'avère néanmoins efficace pour simuler des structures continues présentant une orientation préférentielle et un support composé par des points déjà simulés de taille non négligeable (Fig. 22). Dès que la continuité des objets ne suit pas cette orientation ou que leur sinuosité est importante (chenaux méandriformes), le chemin unilatéral perd en efficacité.

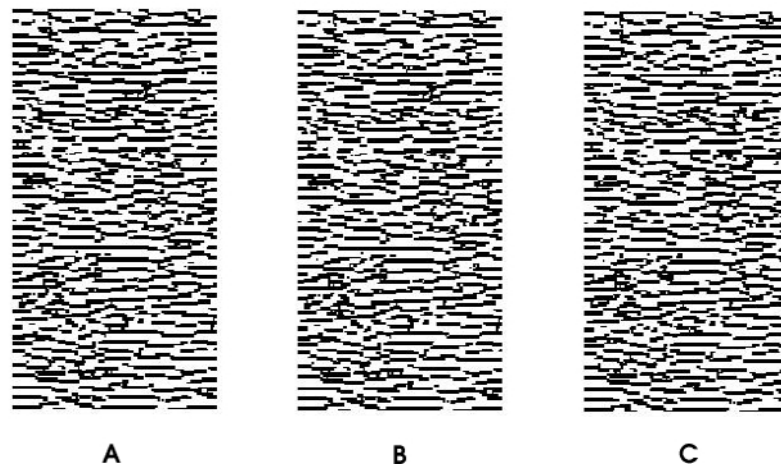


**Fig. 21 Simulation par le chemin régulier (2D).**



**Fig. 22 Simulation MP par le chemin régulier 2D. Reproduction des lignes parallèles : l'image d'apprentissage à droite, la réalisation simulée à gauche.**

Pour raffiner la reproduction de géométries complexes et, par conséquent, la reproduction des objets continus, Remy (2001), Liu et Journal (2004), Strebelle et Remy (2004) proposent d'utiliser le **chemin « post-processus »**. Autrement dit, l'algorithme MP «post-processus» consiste à revenir sur les points simulés avec peu de points voisins et à les resimuler avec plus de points voisins connus. Cette méthode peut être appliquée de manière itérative ou bien par une étape complémentaire en fin de simulation initiale. Nous proposons deux exemples de simulation post-processus d'abord avec 20% des points et puis avec 50% des points à resimuler (Fig. 23). Ces exemples montrent que cette méthode ne garantit pas la reproduction des objets très continus.



**Fig. 23 Simulation MP par le chemin aléatoire post-processus 2D. (A) Reproduction des lignes parallèles : sans resimulation à gauche ; (B) 20% des points resimulés au centre ; (C) 50% des points resimulés à droite. (La simulation est obtenue en utilisant le module SGeMS développé à Stanford).**

Par la suite, nous utiliserons un cheminement aléatoire pour obtenir la plupart des simulations MP sur la grille fine. Nous ne recourons à un cheminement régulier que dans

le cas où la reproduction des objets continus est essentielle.

### 2.3.4 Connexité

La connexité est une caractéristique, qui évalue la continuité (entre autre) des objets géologiques. C'est une caractéristique importante, car elle affecte l'écoulement de fluides dans un réservoir. Lors de la reproduction de la structure interne d'un réservoir, il est important donc de préserver sa connexité. Dans ce paragraphe, nous étudions la qualité de reproduction de la connexité par l'approche MP.

La littérature fournit diverses définitions de la connexité [Krishnan et Journel (2003)]. Dans le cadre de cette étude, nous nous appuyons sur la version classique de la connexité entre deux points [Allard (1994)] :

Déf. Deux points  $u$  et  $u'$  sont proches voisins si leur zone d'influence possède une frontière commune (1 dimension : la frontière est ponctuelle ; 2 dimensions – elle est linéaire ; 3 dimensions – elle est surfacique).

Déf. Deux point  $u$  et  $u'$  sont connectés  $u \leftrightarrow u'$ , s'ils sont reliés par une suite de proches voisins, tous appartenant au même faciès :  $u, u_1, \dots, u_n, u'$ .

Soient  $G$  l'espace discrétisé et  $A$  un ensemble de points dans  $G$  représentant un faciès :  $Z(A) = 1$ ;  $Z(A^c) = 0$ . La fonction de connexité  $\alpha$  entre deux points  $u, u'$  est la probabilité d'avoir ces deux points connectés :

$$\alpha(u, u') = P[u \leftrightarrow u' \mid u, u' \in A]$$

Cette fonction dépend non seulement de la distance entre les deux points  $u$  et  $u'$ , mais aussi de leur position relative. Dans un cadre stationnaire, la fonction de connexité  $\alpha(u, u')$  ne dépend que du vecteur  $u \leftrightarrow u'$ , qui sépare les deux points.

A titre d'illustration, nous avons choisi un cas d'étude basé sur le modèle gaussien seuillé de variogramme hyperbolique. Les exemples sont bidimensionnels et de la taille 100x200. Pour observer les variations de continuité sur nos exemples, nous utilisons plusieurs portées horizontales de variogramme, tandis que la portée verticale reste unitaire. Pour estimer la qualité de reproduction de connexité et de continuité des *clusters* [définition en Annexe B] nous allons comparer les fonctions de covariance et de connexité horizontales dérivées de l'image d'apprentissage et de 10 de ses simulations MP.

La première image d'apprentissage est construite avec un variogramme horizontal de

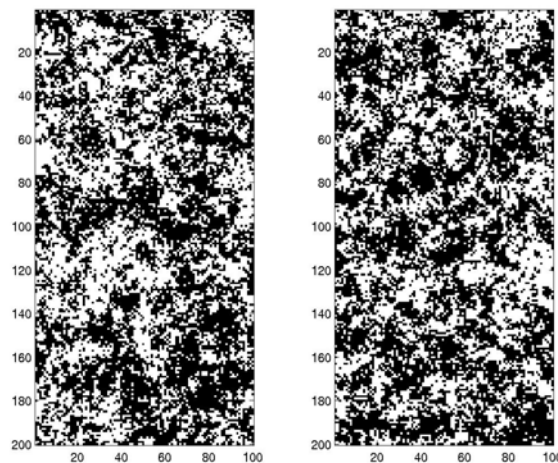


portée unité (Fig. 24). Les fonctions de covariance régionale (Fig. 25) paraissent comparables, ce qui n'est pas vraiment le cas des fonctions de connexité (Fig. 26). Malgré le seuil 0 du modèle gaussien, qui fournit la même proportion entre les deux faciès, la présence de fluctuations statistiques fait que leurs connexités sur l'image d'apprentissage sont assez différentes. Puisque le modèle de simulation MP est fourni par l'image d'apprentissage, l'approche MP reproduit les caractéristiques de l'image d'apprentissage mais pas du modèle gaussien. Cela peut expliquer, pourquoi la connexité des clusters noirs est mieux reproduite.

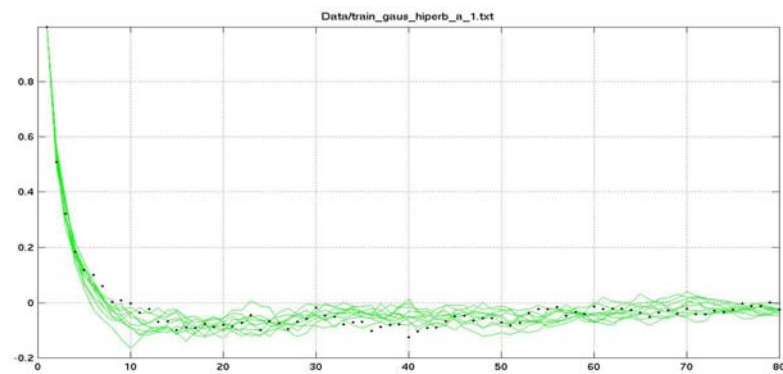
La deuxième image d'apprentissage est construite avec un variogramme horizontal de portée 5 (Fig. 27). Même si l'accord entre les covariances reste satisfaisant (Fig. 28), la connexité des réalisations MP est inférieure à celle de l'image d'apprentissage (Fig. 29).

La troisième et dernière image d'apprentissage est construite avec un variogramme horizontal de portée 20 (Fig. 30). Les covariances montrent, que la portée de la réalisation n'atteint pas 20 (Fig. 31). Cet effet est causé par l'étendue des clusters qui est supérieure non seulement à la taille de fenêtre d'exploration mais aussi à la taille de l'image d'apprentissage. Plus l'étendue (continuité) des clusters augmente, plus les connexités de l'image d'apprentissage et celle des réalisations MP s'avèrent différentes (Fig. 32).

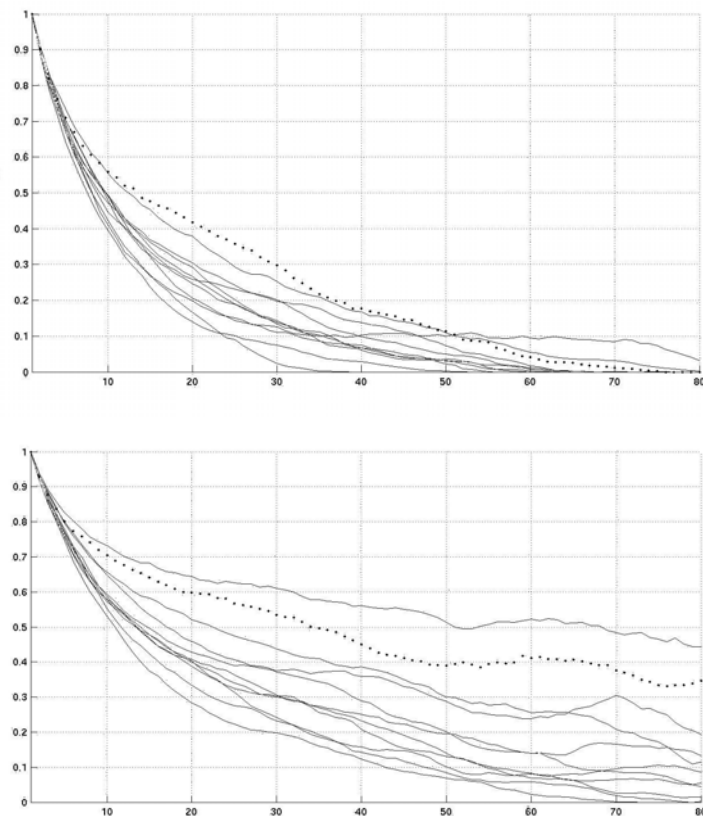
L'exemple précédent nous a montré que les multiples réalisations MP peuvent présenter des connexités assez variées. Ainsi, la connexité peut jouer un rôle important dans le choix d'une réalisation par rapport aux autres. Dans l'exemple suivant, nous allons montrer, que la fluctuation de connexité peut indiquer la qualité de reproduction de la géométrie et donc, doit être prise en compte comme caractéristique qualificative.



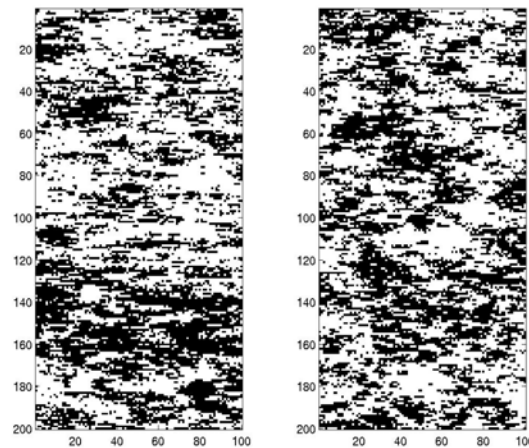
**Fig. 24 Simulation stationnaire par l'approche MP. Reproduction du modèle gaussien hyperbolique seuillé (portée horizontale unité) : l'image d'apprentissage à gauche, l'une des 10 réalisations MP à droite.**



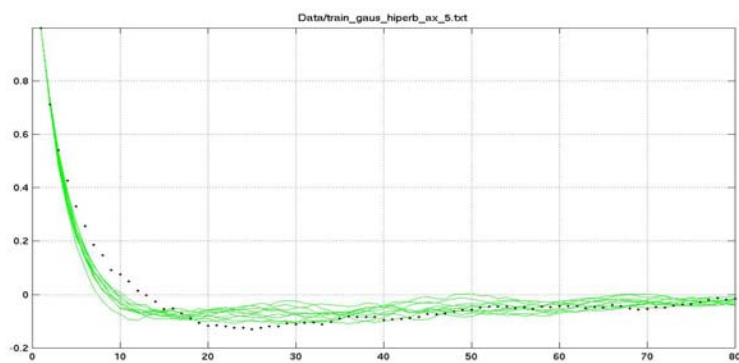
**Fig. 25** Fonction de covariance horizontale du modèle gaussien hyperbolique seuillé (portée horizontale unité) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.



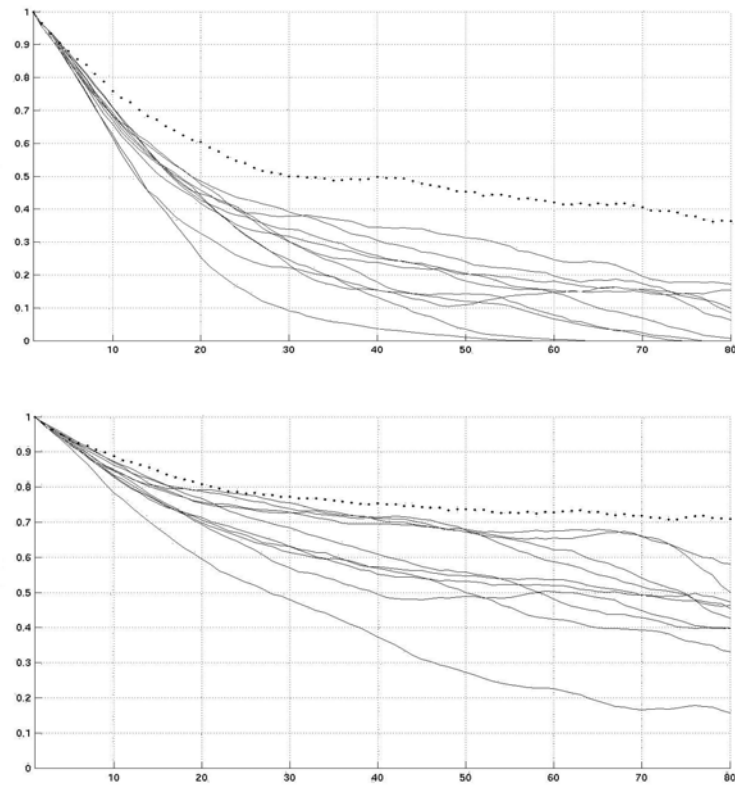
**Fig. 26** Fonction de connexité horizontale du modèle gaussien hyperbolique seuillé (portée horizontale unité) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP ; en haut pour les clusters noirs, en bas pour les clusters blancs.



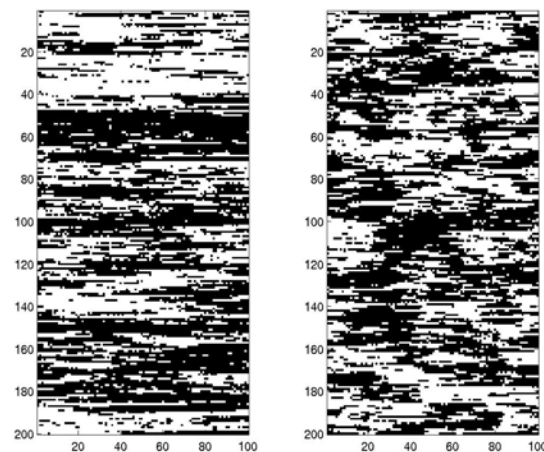
**Fig. 27** Simulation stationnaire par l'approche MP. Reproduction du modèle gaussien hyperbolique seuillé (portée horizontale égale à 5) : l'image d'apprentissage à gauche, une des 10 réalisations MP à droite.



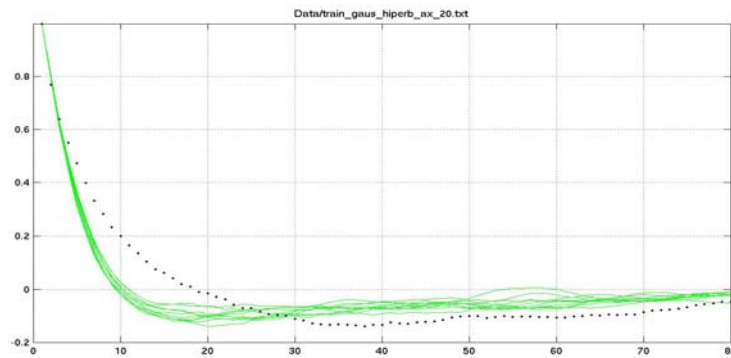
**Fig. 28** Fonction de covariance horizontale du modèle gaussien hyperbolique seuillé (portée horizontale égale à 5) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.



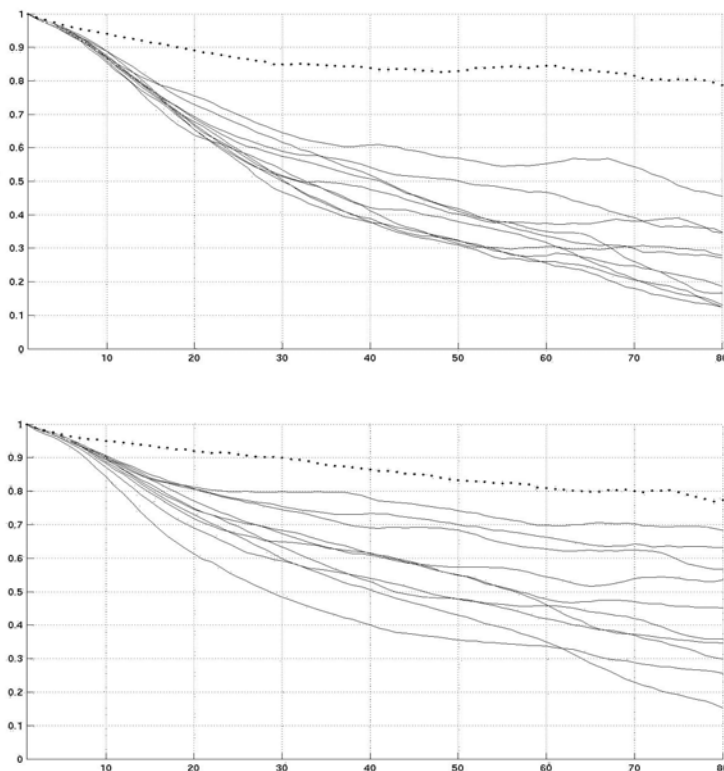
**Fig. 29** Fonction de connexité horizontale du modèle gaussien hyperbolique seuillé (portée horizontale égale à 5) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP ; en haut pour les clusters noirs, en bas pour les clusters blancs.



**Fig. 30** Simulation stationnaire par l'approche MP. Reproduction du modèle gaussien hyperbolique seuillé (portée horizontale égale à 20) : l'image d'apprentissage à gauche, une des 10 réalisations MP à droite.



**Fig. 31** Fonction de covariance du modèle gaussien hyperbolique seuillé (portée horizontale égale à 20) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.



**Fig. 32** Fonction de covariance du modèle gaussien hyperbolique seuillé (portée horizontale égale à 20) : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP; en haut pour les clusters noirs, en bas pour les clusters blancs.

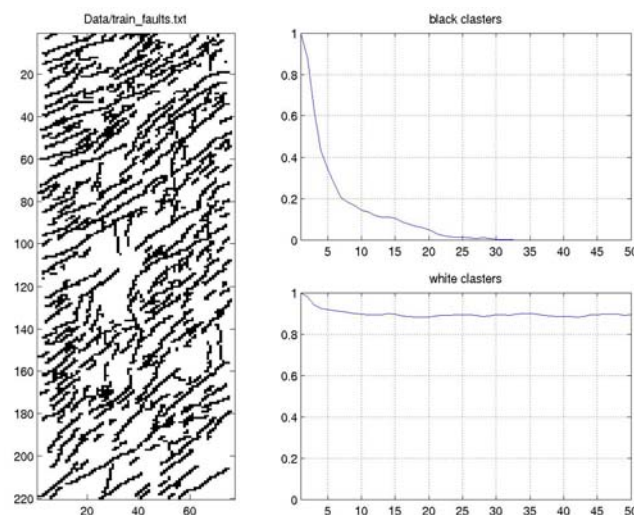
Le cas d'étude suivant est basé sur un réseau de fractures d'un gisement situé en Afrique du Nord et qui a déjà fait l'objet de plusieurs études [Gauthier et al. (2002)]. Nous disposons d'une image d'apprentissage en 2D de taille 100x200 représentant le réseau de fractures à grande échelle (Fig. 33). Dix réalisations sont simulées à partir de l'image d'apprentissage par approche MP (Fig. 34 - 35). Dix fonctions de connexité sont extraites de

ces réalisations et comparées avec celle de l'image d'apprentissage (Fig. 36).

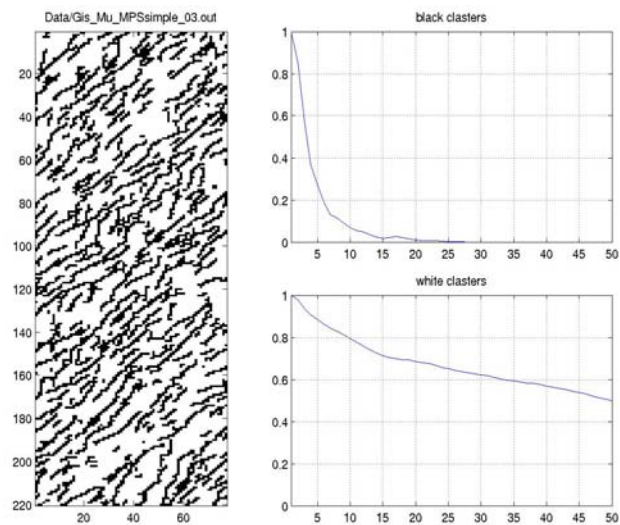
Toutes les réalisations reproduisent assez bien la connexité verticale des fractures, pourtant elles ne parviennent pas à reproduire correctement celle de leur complémentaires (clusters blancs). Nous pouvons également observer que les fractures présentent une connexité très peu de fluctuante, au contraire de leur complémentaire. Cela résulte vraisemblablement de ce que les fractures sont mieux reproduites que leur complémentaire. En revenant au paragraphe 2.3.2, ce défaut s'explique par l'allongement des clusters blancs qui est comparable à celle de l'image d'apprentissage ; ils sont donc supérieurs à la taille de fenêtre d'exploration.

A partir des résultats de cette étude, nous tirons les conclusions suivantes :

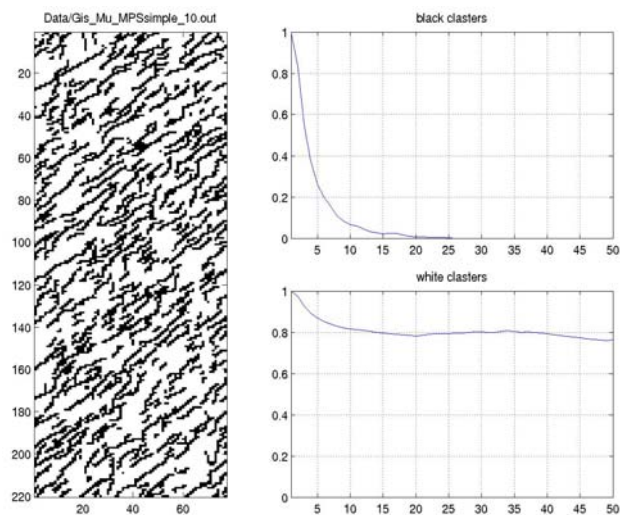
- 1) La connexité peut être reproduite par l'approche MP si les paramètres de simulation sont adaptés à la géométrie.
- 2) La connexité peut être à la fois une caractéristique quantitative (continuité des faciès) et qualitative de la simulation MP :
  - faible fluctuation et l'accord avec la connexité de l'image d'apprentissage nous assure la bonne reproduction de la géométrie y compris la continuité des faciès ;
  - fluctuation importante et désaccord avec la connexité de référence nous signale que la reproduction de la géométrie est mauvaise et que les paramètres de simulation ne sont pas adaptés à la géométrie.



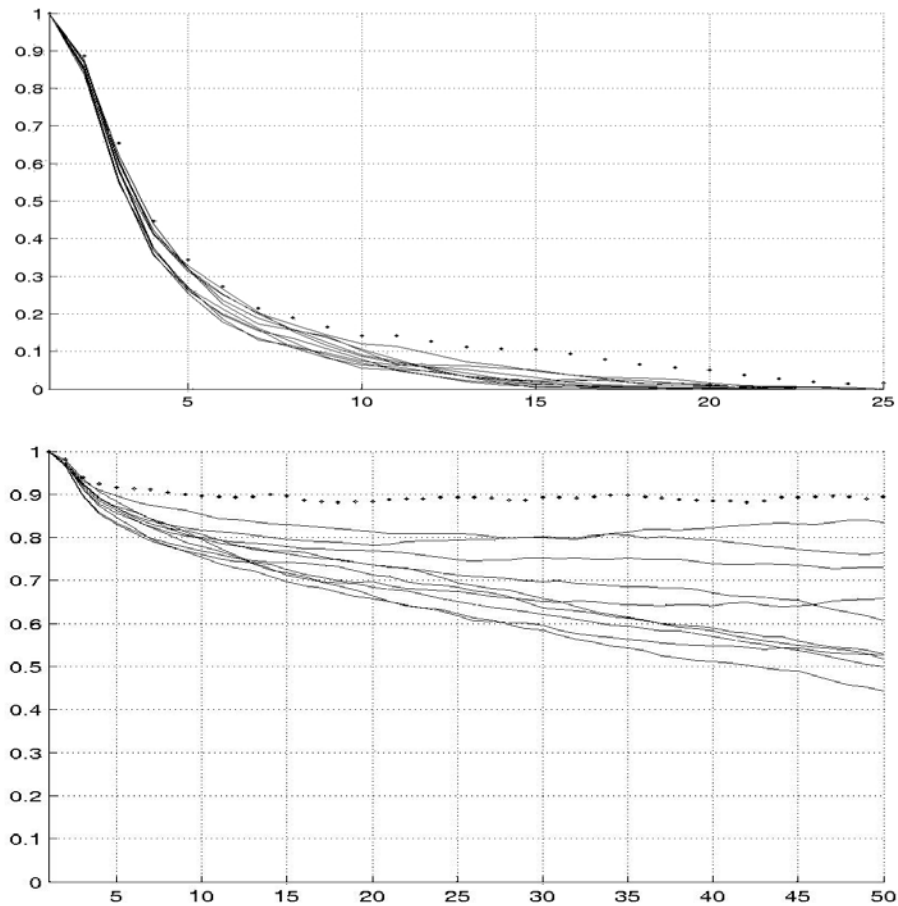
**Fig. 33 L'image d'apprentissage du réseau de fractures. Fonction de connexité verticale : en haut pour les clusters noirs, en bas pour les clusters blancs.**



**Fig. 34** Une des 10 réalisations MP. Fonction de connexité verticale : en haut pour les clusters noirs, en bas pour les clusters blancs.



**Fig. 35** Une des 10 réalisations MP. Fonction de connexité verticale : en haut pour les clusters noirs, en bas pour les clusters blancs.



**Fig. 36** Fonction de connexité verticale du réseau de fractures et ses 10 réalisations MP : la ligne pointillée correspond à l'image d'apprentissage et les lignes pleines aux 10 réalisations MP.

### 2.3.5 Stationnarité

L'utilisation classique de l'approche MP nécessite la stationnarité spatiale pour pouvoir construire une loi empirique à partir des statistiques acquises de l'image d'apprentissage quelque soit leur localisation spatiale. Il existe plusieurs formes de stationnarité :

#### 1) stationnarité forte :

Rappelons que la fonction aléatoire  $Z(\cdot)$  est stationnaire si pour n'importe quel nombre de points  $u_1, u_2, \dots, u_n$  et n'importe quel  $h$  la distribution conjointe des fonctions aléatoires  $Z(u_i)$ ,  $i = 1, \dots, n$  est égale à la distribution conjointe des



fonctions aléatoires  $Z(u_i + h)$ ,  $i = 1, \dots, n$  :

$$F\{Z(u_1), Z(u_2), \dots, Z(u_n)\} = F\{Z(u_1 + h), Z(u_2 + h), \dots, Z(u_n + h)\}.$$

Ce type de stationnarité est difficilement vérifiable et rarement utilisé dans la pratique géostatistique.

- 2) **stationnarité faible** ou stationnarité du second ordre. Toutes les variables ont la même moyenne. De plus, la covariance entre deux points d'appui ne dépend pas spécifiquement d'eux, mais seulement du vecteur qui les sépare. Cette forme de stationnarité beaucoup plus lâche n'est toutefois pas toujours satisfaite.
- 3) **hypothèse intrinsèque**. Elle ne nécessite que la stationnarité de la différence de premier ordre  $Z(u_i + h) - Z(u_i)$ ,  $i = 1, \dots, n$ . Cette forme de stationnarité est très souvent utilisée en géostatistique et en analyse des séries temporelles.
- 4) **stationnarité faible avec dérive**. L'idée de cette forme de stationnarité est qu'on peut se débarrasser de la dérivée pour se retrouver dans un contexte de stationnarité faible.

L'approche MP a été développée sous l'hypothèse de stationnarité forte de la fonction aléatoire  $Z$ .

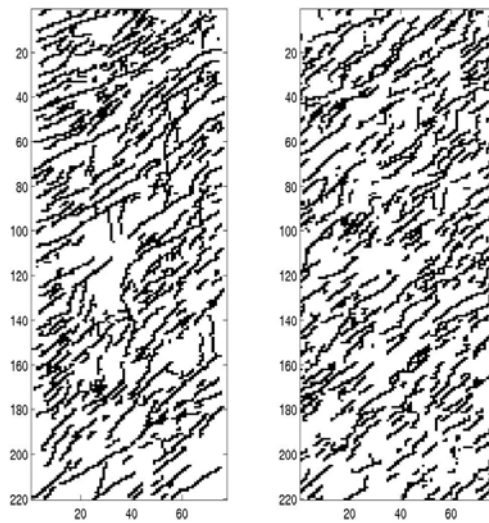
Précisons que la stationnarité est une caractéristique d'une fonction aléatoire  $Z(\cdot)$  et non pas des données spatiales que l'on cherche à modéliser. La conséquence est que la forme de stationnarité de  $Z$  ne peut pas être strictement inférée à partir d'une ou de plusieurs réalisations de cette fonction aléatoire. Elle peut par contre être présumée. Il est possible de tester après coup la légitimité de notre hypothèse en étudiant la compatibilité des réalisations de  $Z(\cdot)$  à nos données. Il est à noter que ce test ne permet que de réfuter l'hypothèse faite. La non réfutation de l'hypothèse ne garantit pas la forme de stationnarité retenue, mais laisse cette éventualité possible.

En pratique, le cas stationnaire correspond au cas où l'image d'apprentissage est suffisamment homogène pour que les statistiques que l'on en extrait, puissent être considérées comme représentatives de toute l'image. La vérification pratique de l'hypothèse de stationnarité à partir d'une image souligne l'importance d'un facteur d'échelle qui est totalement absent du formalisme probabiliste [Chauvet (1999)]. Un même objet géométrique peut être considéré comme constant ou variable, structuré ou non, selon l'échelle à laquelle il sera examiné.

Prenons l'exemple du réseau de fractures (Fig. 37). Supposons, que notre image d'apprentissage est une réalisation d'un processus stationnaire. Cette hypothèse est plausible si, par exemple, on travaille à l'échelle d'un groupe de fractures. Une simulation

MP a été réalisée sous hypothèse de stationnarité (Fig. 37). Il est intéressant de voir que la répartition des fractures de la réalisation est plus homogène que celle de l'image d'apprentissage. Si l'on travaille à l'échelle d'une fracture, la variabilité de taille, d'orientation et de présence de faciès est enregistrée.

Cet exemple illustre parfaitement le fait que la stationnarité est une notion dépendante de l'échelle de travail. À la question de savoir si la réalisation donnée peut être considérée comme la réalisation d'une fonction stationnaire, la réponse dépend finalement de l'objectif de modélisation (ce qui est d'ailleurs le cas pour toute modélisation géostatistique) et des caractéristiques que l'on veut préserver par cette modélisation.



**Fig. 37 Simulation stationnaire par l'approche MP. Reproduction d'un réseau de fractures : l'image d'apprentissage (à gauche) et la simulation MP (à droite) avec la proportion de faciès respectivement égale à 0.28 et 0.29.**

### 2.3.6 Résumé

Pour conclure la partie des conditions d'application de l'approche MP, nous rappelons les 4 conditions nécessaires pour une bonne reproduction d'une structure géométrique :

- 1) L'image d'apprentissage doit contenir les patterns géométriques représentatifs de la structure et répétitifs dans l'espace (de faible entropie) ;
- 2) La taille de l'image d'apprentissage doit être suffisamment grande par rapport à la taille de fenêtre d'exploration ;
- 3) La taille de la fenêtre d'exploration doit être suffisamment grande par rapport à la taille des patterns répétitifs de l'image d'apprentissage ;

- 4) La fonction aléatoire  $Z(\cdot)$  est présumée stationnaire. De plus, la variabilité spatiale de l'hétérogénéité de l'image d'apprentissage doit se stabiliser au-delà d'une certaine distance bien inférieure à la taille de l'image.

Si ces 4 conditions d'applications sont respectées, l'application de l'approche MP garantit une reproduction correcte de la géométrie en préservant ses caractéristiques (formes, orientation, continuité ...). Dans ce cas, le chemin de simulation peut être aléatoire, ce qui permet une meilleure reproduction de la diversité des patterns présentés dans l'image d'apprentissage. Dès que l'une de ces conditions n'est pas respectée, l'application de l'approche MP ne garantit pas une reproduction correcte de la géométrie. Dans ce cas, un ajustement des paramètres peut améliorer les résultats de simulation MP :

- 1) La configuration et la taille de la fenêtre d'exploration sont choisies en fonction de la géométrie de patterns répétitifs ;
- 2) Le cheminement régulier de simulation est défini de manière à pouvoir suivre l'orientation préférentielle des structures continues ;
- 3) Une simulation multigrille est appliquée pour reproduire les objets géométriques de taille supérieure à celle de fenêtre d'exploration.

Après avoir obtenu une réalisation MP, une vérification de sa qualité de reproduction est conseillée. Les caractéristiques d'image (moyenne, écart-type ...) peuvent faire apparaître les discordances majeures entre l'image d'apprentissage et sa réalisation MP. Les moyennes locales et les statistiques multipoints permettent de comparer plus finement leur géométrie. De son côté, la fonction de connexité aide à vérifier toute la conformité de la continuité et aussi de la géométrie dans les deux images.

## 2.4 Vers la simulation non stationnaire

Nous venons de voir le fonctionnement de l'approche multipoints moyennant l'introduction implicite d'une hypothèse de stationnarité (§2.3.5). En pratique toutefois, la réalité est souvent non stationnaire. Les structures géométriques formées par l'hétérogénéité du réservoir possèdent des caractéristiques (dites morphologiques) variables dans l'espace. Voici plusieurs caractéristiques d'un réservoir pétrolier susceptibles d'évoluer de façon non stationnaire :

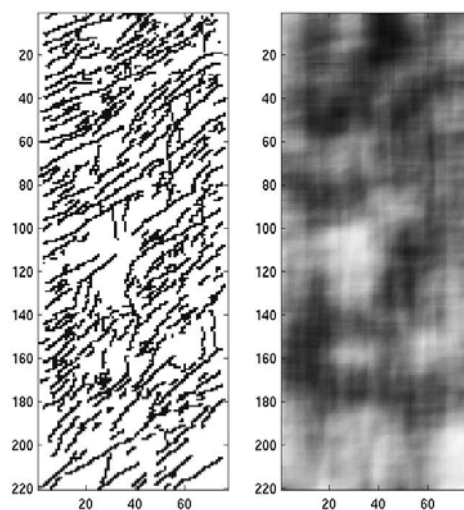
- proportion locale (densité) de faciès ;
- taille des objets géologiques ;
- orientation régionale des objets ;
- connexité régionale ;
- etc.

Si la variabilité spatiale de l'hétérogénéité de l'image d'apprentissage ne se stabilise au-delà d'une certaine distance bien inférieure à la taille de cette image, nous la considérons comme non stationnaire. Dans ce cas, il est intéressant de préserver et de reproduire la non stationnarité spatiale par simulation MP.

Cette non stationnarité est souvent accessible via une information, issue, par exemple, de la sismique [Gauthier et al. (2002)]. La Fig. 38 montre un exemple dans un cadre non stationnaire spatiale sur le réseau de fractures qui l'on a déjà utilisé dans un cadre stationnaire (§§2.3.4-2.3.5). L'image d'apprentissage présente des zones de faible et de forte fracturation, ce qui est confirmé par la carte sismique.

Une autre source de données non stationnaires est l'image d'apprentissage elle-même. Dans ce cas, les caractéristiques géométriques de forte variabilité spatiale peuvent être extraites de l'image par des méthodes morphologiques tel que la moyenne mobile ou l'analyse de clusters ...).

Dans le chapitre suivant, nous allons montrer comment prendre en compte ces données dans une simulation multipoints.



**Fig. 38** Image d'apprentissage à gauche ; sa carte de densité de fracturation à droite.

### 3 Simulation non stationnaire : approches existantes

Jusqu'à présent, l'approche géostatistique multipoints a été développée dans un cadre stationnaire, la loi spatiale à simuler étant invariante par la translation. Plaçons nous maintenant dans le cas où nous disposons d'une information non stationnaire. La littérature fournit deux méthodes pour prendre en compte cette information dite auxiliaire dans une simulation multipoints.

La plus populaire est celle proposée par Journel (2002) et ultérieurement développée par Krishnan et al. (2004). C'est un modèle analytique connu sous le nom de « modèle Tau ». Il s'agit d'une approximation de la probabilité biconditionnelle en fonction de probabilités monoconditionnelles avec un paramètre  $\tau$  à ajuster au cas par cas. Ce paramètre définit l'impact des données auxiliaires sur le résultat final de simulation. Dans la pratique usuelle de cette méthode,  $\tau$  est pris égale à 1, à défaut de pouvoir faire mieux. Ce paramètre est difficilement évaluable [Krishnan (2004)] ; il dépend non seulement de la situation particulière étudiée mais aussi du point à simuler.

L'autre méthode prend appui sur une méthode empirique développée par Strebelle (2002), Strebelle et Zhang (2004). L'idée de cette méthode est d'extraire la probabilité conditionnelle directement de l'image d'apprentissage. Pour ce faire, les données auxiliaires sont réparties en classes, ce que réduit leur variabilité et limite leur représentativité.

### 3.1 Description du problème

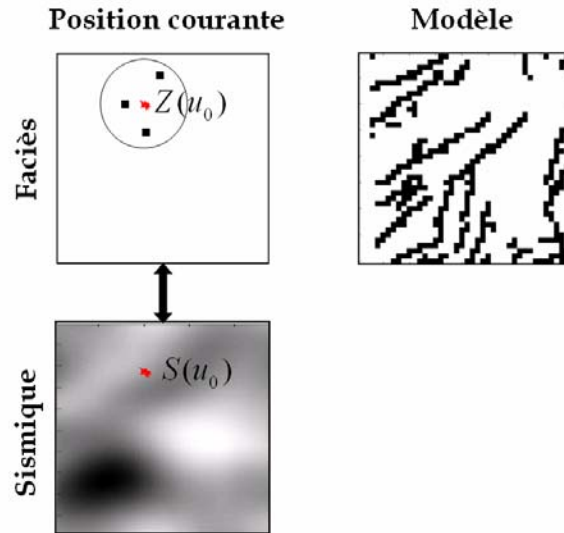


Fig. 39 Espace de simulation avec sa carte de contrainte spatiale à gauche ; image d'apprentissage à droite.

Revenons à l'algorithme MP en ajoutant une fonction auxiliaire. Dans le cas illustré par notre exemple (Fig. 39), c'est la proportion locale de faciès. Elle nous indique les zones de forte et de faible apparition des différents faciès (dans le cas général, cette fonction pourra représenter des caractéristiques non stationnaires très diverses).

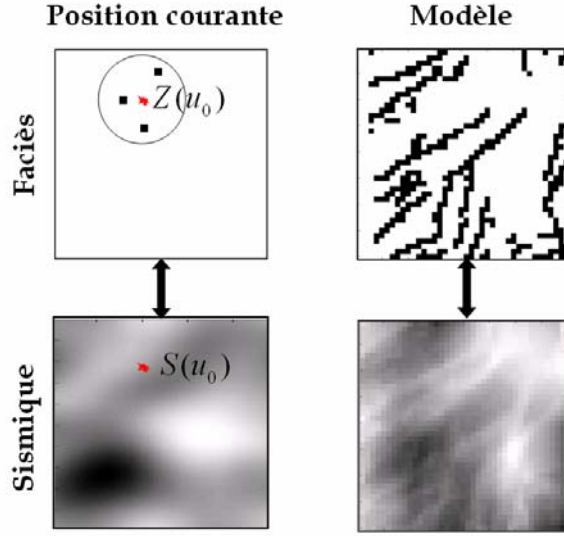
Soit  $S$  une fonction auxiliaire associée à la fonction aléatoire principale  $Z$  et définie  $\forall u \in E$ . Dans le cas général, les valeurs prises par la fonction auxiliaire ne sont pas limitées au nombre de faciès de sorte que l'on peut écrire  $S(u) \in R$ ,  $\forall u \in E$ .

Considérons un point  $u_0$  à simuler. La simulation passe par l'évaluation de la loi de  $Z(u_0)$ , conditionnée par les valeurs déjà simulées dans son voisinage  $Z(\mathbf{u}) = \mathbf{z}$  ainsi que par la valeur de la fonction auxiliaire  $S$  en  $u_0$ :

$$P[Z(u_0) | Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s] = ?$$

## 3.2 Méthode de classement

### 3.2.1 Principe



**Fig. 40** Espace de simulation avec sa carte de contrainte spatiale à gauche ; images d'apprentissage principale et auxiliaire à droite.

La méthode de classement consiste à utiliser, en plus d'une image d'apprentissage de la fonction principale, une image d'apprentissage de la fonction auxiliaire (Fig. 40). (Cette dernière est obtenue, par exemple, par simulation numérique à partir de la première.) Étant donnés les états du voisinage de la fonction principale et de la valeur de la fonction auxiliaire en  $u_0$ , nous identifions l'ensemble des configurations des points compatibles à la fois avec l'état de ce voisinage et avec cette valeur auxiliaire :

$$W = \{w_0 \in \mathbf{I} : Z(\mathbf{w}) = \mathbf{z}, \mathbf{w} = \mathbf{u}, S(w_0) = s\}$$

Nous formons ensuite la loi empirique conditionnelle de  $Z(u_0)$  :

$$\begin{aligned} P[Z(u_0) = k \mid Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s] &= \frac{P[Z(u_0) = k, Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s]}{P[Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s]} \\ &= \frac{\#\{w_0 \in W : Z(w_0) = k\}}{\#\{w_0 \in W\}}, \quad k = 1, 2, \dots, K \end{aligned}$$

Cette méthode directe et cohérente présente toutefois un inconvénient. En raison du caractère souvent continu de la fonction auxiliaire, il y a généralement très peu de chance de trouver une configuration, dans les deux images d'apprentissage, qui soit compatible aux deux conditions. Une façon de tourner cette difficulté est de deviser les valeurs de la fonction auxiliaire en plusieurs classes (Fig. 41). Cela atténue le problème mais introduit un



certain arbitraire dans l'algorithme proposé :

$$P[Z(u_0) = k \mid Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s] \approx \frac{P[Z(u_0) = k, Z(\mathbf{u}) = \mathbf{z}, s_i < S(u_0) < s_{i+1}]}{P[Z(\mathbf{u}) = \mathbf{z}, s_i < S(u_0) < s_{i+1}]}$$

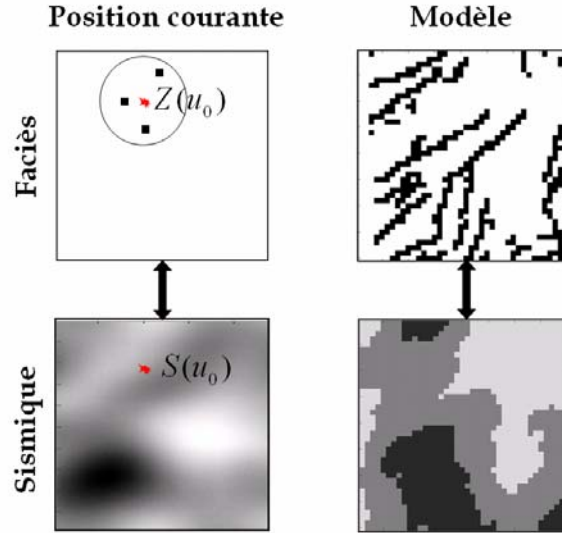


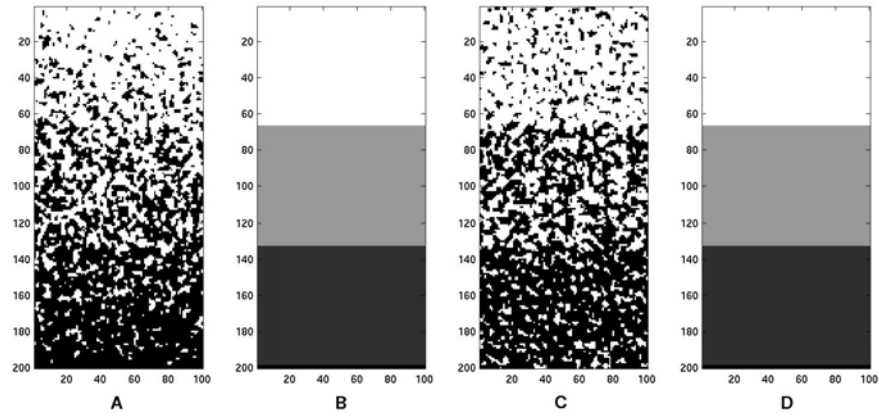
Fig. 41 Espace de simulation avec sa carte de contrainte spatiale à gauche ; images d'apprentissage principale et auxiliaire classée à droite.

### 3.2.2 Application

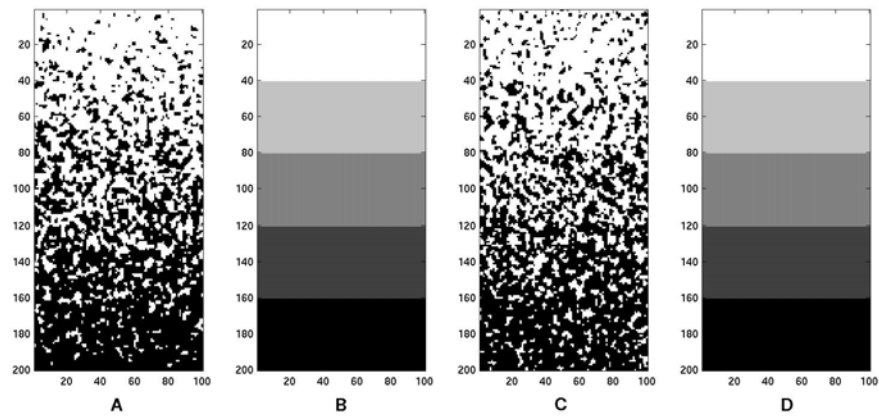
Voici un exemple d'application de la méthode de classement.

L'image d'apprentissage principale est une réalisation d'une fonction gaussienne 2D de variogramme gaussien (portée 3) avec une dérive verticale. La fonction est seuillée en 2 faciès de telle manière que la proportion locale de faciès varie continûment de 1 à 0 le long de la verticale. Dans notre cas d'étude, cette proportion définit l'image d'apprentissage auxiliaire de même que la contrainte spatiale.

L'application de la méthode de classement consiste à discrétiser la fonction auxiliaire continue en un nombre limité des classes. La première simulation est réalisée avec le classement en 3 intervalles  $[0, 0.33]$ ;  $[0.33, 0.66]$ ;  $[0.66, 1.0]$ ; il en résulte une réalisation MP en trois parties clairement identifiables (Fig. 42). En augmentant le nombre de classes (5 intervalles  $[0, 0.2]$ ;  $[0.2, 0.4]$ ;  $[0.4, 0.6]$ ;  $[0.6, 0.8]$ ;  $[0.8, 1.0]$ ), nous diminuons légèrement cet artefact (Fig. 43).



**Fig. 42** (A) Image d'apprentissage principale, (B) image d'apprentissage auxiliaire discrétisée en 3 classes, (C) réalisation simulée selon (D) sa carte de contrainte spatiale.



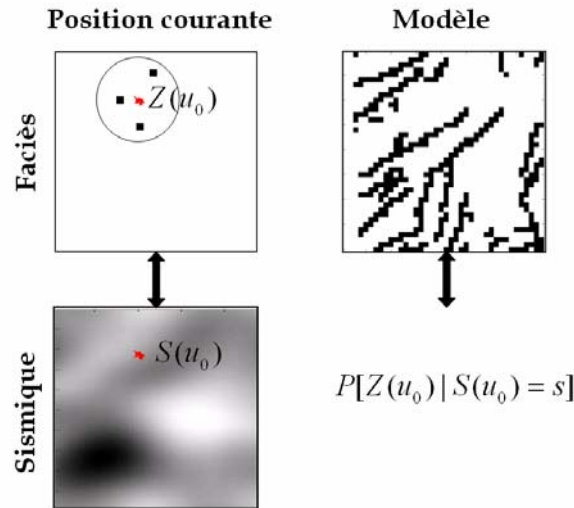
**Fig. 43** (A) Image d'apprentissage principale, (B) image d'apprentissage auxiliaire discrétisée en 5 classes, (C) réalisation simulée selon (D) sa carte de contrainte spatiale.

### 3.2.3 Remarques

Nous venons de voir le fonctionnement de l'approche MP conditionnée à une contrainte spatiale non stationnaire à l'aide de la méthode de classement. Cette approche nous semble prometteuse malgré l'artefact de simulation causé par le regroupement en classes des données auxiliaires. Le fait d'utiliser deux images d'apprentissage permet de reproduire la structure géométrique semblable à celle de l'image d'apprentissage conformément à la contrainte spatiale donnée. De plus, cette méthode est indépendante de la nature des contraintes spatiales, ce qui n'est pas le cas pour le modèle Tau présenté ci-dessous.

## 3.3 Modèle Tau

### 3.3.1 Principe



**Fig. 44** Espace de simulation avec sa carte de contrainte spatiale à gauche ; image d'apprentissage à droite.

Pour préserver la présentation originale de cette approche, nous conservons les notations proposées par Journal (2002).

Soit  $A$  l'événement qui est la présence du faciès  $k$  au point  $u_0$  :

$$A: \{Z(u_0) = k\}$$

Soit  $B$  l'événement qui donne l'état du voisinage du point  $u_0$  :

$$B: \{Z(\mathbf{u}) = \mathbf{z}\}$$

Soit  $C$  l'événement qui donne la valeur prise par la fonction auxiliaire au point  $u_0$  :

$$C: \{S(u_0) = s\}$$

Avec ces notations, nous cherchons à estimer la probabilité  $P[A|B, C]$ . Supposons que l'on connaisse les probabilités suivantes :  $P[A]$ ,  $P[A|B]$ ,  $P[A|C]$ . L'idée du modèle Tau est d'estimer la probabilité biconditionnelle à partir de ces probabilités monoconditionnelles :

$$P[A|B, C] = f_\tau(P[A], P[A|B], P[A|C])$$

Pour cela, Journal fait l'hypothèse suivante :

« La contribution incrémentale de l'événement  $C$  sur la connaissance de l'événement  $A$  est la même avant ou après la connaissance de l'événement  $B$ . »

Définissons les paramètres  $a, b, c, x$  comme les *distances logiques* des événements  $\{A\}, \{A|B\}, \{A|C\}$  et  $\{A|B,C\}$ . Ces distances augmentent en raison inverse de la probabilité des événements correspondants. Par exemple, pour l'événement  $A$  nous avons :

$$a = \frac{1 - P[A]}{P[A]}; \quad P[A]=1 \Leftrightarrow a=0; \quad P[A]=0 \Leftrightarrow a=\infty.$$

En termes de distance logique, l'hypothèse de Journal s'écrit comme une relation de proportion :

$$\frac{x}{b} = \frac{c}{a}$$

où

$$a = \frac{1 - P[A]}{P[A]}, \quad b = \frac{1 - P[A|B]}{P[A|B]}, \quad c = \frac{1 - P[A|C]}{P[A|C]}, \quad x = \frac{1 - P[A|B,C]}{P[A|B,C]}.$$

Pour pouvoir mieux contrôler l'impact de l'événements  $C$ , l'hypothèse de Journal est enrichie par l'introduction d'un paramètre  $\tau$ , idéalement ne dépendant que de  $B$  et de  $C$ , et tel que :

$$\frac{x}{b} = \left( \frac{c}{a} \right)^{\tau(B,C)}$$

Ainsi, la probabilité recherchée peut être approximée par une fonction des probabilités connues :

$$P^*[A|B,C] = \frac{1}{1 + \left( \frac{c}{a} \right)^{\tau} b} \quad 3.1$$

Krishnan et al. (2004) interprète ce modèle comme une expression exacte de la probabilité biconditionnelle, car le paramètre  $\tau$  peut être évalué explicitement :

$$\tau = \log \left( \frac{P[C|A^c,B]}{P[C|A,B]} \right) \bigg/ \log \left( \frac{P[C|A^c]}{P[C|A]} \right) \quad 3.2$$

Ceci n'est possible que si l'on connaît les probabilités conditionnelles  $P[C|A,B]$ ,  $P[C|A^c,B]$ ,  $P[C|A]$ ,  $P[C|A^c]$ . Sous cette condition, nous n'avons plus besoin du modèle Tau pour calculer la valeur explicite de la probabilité biconditionnelle :

$$P[A|B,C] = \frac{1}{1 + \frac{P[C|A^c,B]}{P[C|A,B]} \cdot \frac{P[A^c|B]}{P[A|B]}}$$

Puisque ni  $P[C|A,B]$ ,  $P[C|A^c,B]$ ,  $P[C|A]$ ,  $P[C|A^c]$ , ni  $P[A|B,C]$  sont pas connues, l'inférence exacte de paramètre  $\tau$  (Eq. 3.2) est impossible. En pratique, on attribue souvent à  $\tau$  la valeur 1. Le modèle Tau reste donc une approximation de probabilité

biconditionnelle.

Ce modèle n'a de sens que si le paramètre  $\tau$  ne dépend pas de l'événement  $A$ .

### 3.3.2 Validation sur trois modèles stochastiques

Dans le cadre de cette étude, le modèle Tau a été testé sur trois modèles probabilistes :

- 1) modèle trigaussien ;
- 2) modèle trigaussien seuillé ;
- 3) modèle trivariable autodual.

Le but de cette validation est de comparer les différentes probabilités biconditionnelles avec leur approximation par le modèle Tau. Le choix des modèles s'explique par le fait que chacun d'eux a ses probabilités jointes (aussi conditionnelles et biconditionnelles) calculables de manière explicite.

**Modèle trigaussien.** Soit  $(X, Y, Z)$  un vecteur gaussien standard de matrice de covariance  $\Lambda$  :

$$\Lambda = \begin{pmatrix} 1 & \rho_{XY} & \rho_{XZ} \\ \rho_{XY} & 1 & \rho_{YZ} \\ \rho_{XZ} & \rho_{YZ} & 1 \end{pmatrix}.$$

Considérons trois événements aléatoires définis comme tels :  $A = \{X < x\}$  ;  $B = \{Y = y\}$  ;  $C = \{Z = z\}$ . Les probabilités nécessaires à la spécification du modèle Tau s'écrivent :

$$\begin{aligned} P[A] &= P[X < x] ; \\ P[A|B] &= P[X < x | Y = y] ; \\ P[A|C] &= P[X < x | Z = z] ; \\ P[A|B, C] &= P[X < x | Y = y, Z = z]. \end{aligned}$$

Soit  $g(\cdot)$  la densité du vecteur  $(X, Y, Z)$ . Alors, les densités de  $X$  sachant  $Y = y$  et sachant  $Z = z$  valent respectivement :

$$\begin{aligned} g(x|y) &= \frac{g(x, y)}{g(y)} = \frac{1}{2\pi\sqrt{1-\rho_{XY}^2}} e^{-\frac{1}{2(1-\rho_{XY}^2)}(x-\rho_{XY}y)^2} ; \\ g(x|z) &= \frac{g(x, z)}{g(z)} = \frac{1}{2\pi\sqrt{1-\rho_{XZ}^2}} e^{-\frac{1}{2(1-\rho_{XZ}^2)}(x-\rho_{XZ}z)^2} . \end{aligned}$$

De son coté, la densité du  $X$  sachant  $Y = y$  et  $Z = z$  s'écrit :

$$g(x | y, z) = \frac{g(x, y, z)}{g(y, z)} = \frac{\sqrt{1 - \rho_{YZ}^2}}{\sqrt{2\pi} \|\Lambda\|} e^{-\frac{1 - \rho_{YZ}^2}{2\|\Lambda\|} \left( x + \frac{\rho_{XZ}\rho_{YZ} - \rho_{XY}}{1 - \rho_{YZ}^2} y + \frac{\rho_{XY}\rho_{YZ} - \rho_{XZ}}{1 - \rho_{YZ}^2} z \right)^2}$$

On désignera par  $G(x|y)$ ,  $G(x|z)$  et  $G(x|y, z)$  les fonctions de répartition correspondantes. Les distances logiques s'expriment à l'aide de ces fonctions de répartition conditionnelles :

$$a = \frac{1 - G(x)}{G(x)}, \quad b = \frac{1 - G(x|y)}{G(x|y)}, \quad c = \frac{1 - G(x|z)}{G(x|z)} \quad 3.3$$

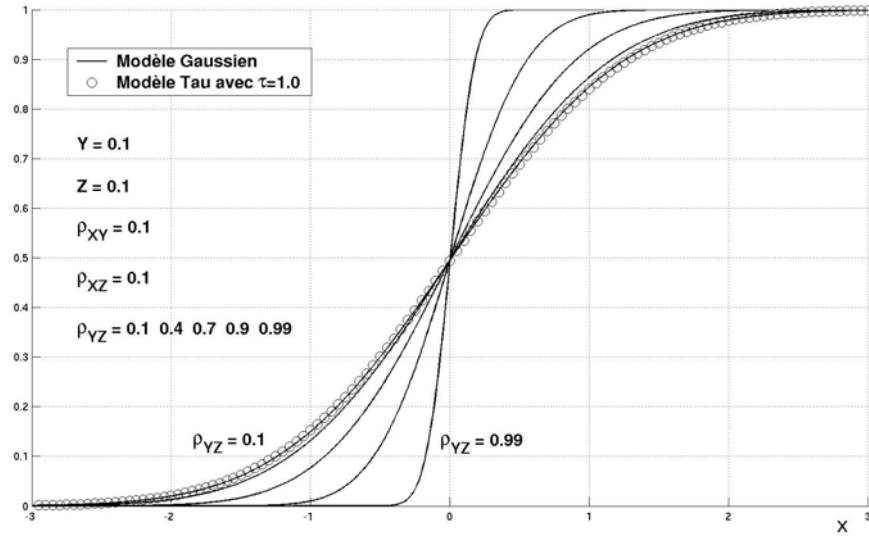
Une fois ces distances calculées, nous obtenons une approximation de la probabilité biconditionnelle (Eq. 3.1).

Évaluons tout d'abord le degré d'approximation de la probabilité biconditionnelle par le modèle Tau avec  $\tau = 1$ . Ainsi, nous comparons deux modèles qui ne sont fonctions que de la variable  $x$  une fois que  $y$  et  $z$  ont été fixées :

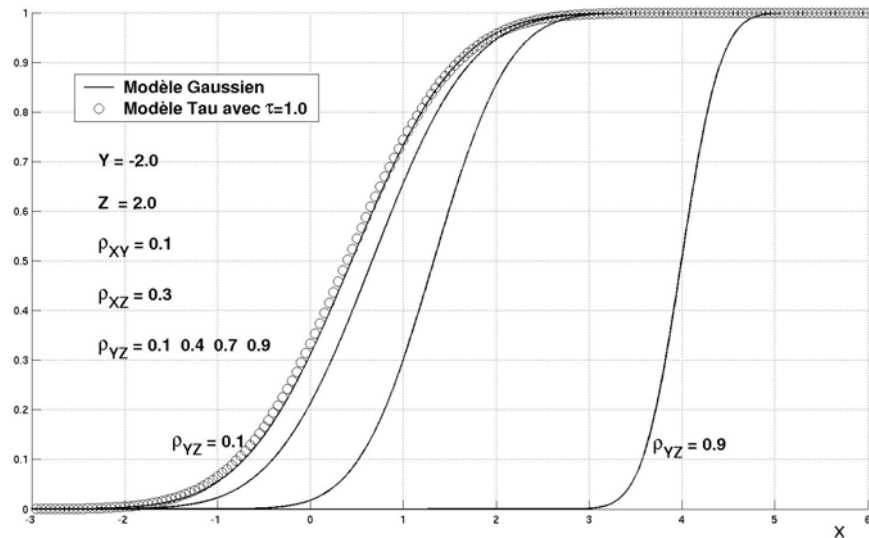
$$P[A | B, C] = G(x | y, z)$$

$$P^*[A | B, C] = \frac{1}{1 + \left(\frac{c}{a}\right)^\tau b} \quad 3.4$$

Un premier test est effectué avec des valeurs gaussiennes proches de la moyenne  $y = 0.1$ ,  $z = 0.1$  et avec des coefficients de corrélation faibles  $\rho_{XY} = 0.1$  et  $\rho_{XZ} = 0.1$  (Fig. 45). Comme deuxième test nous avons pris des valeurs gaussiennes  $y = -2.0$ ,  $z = 2.0$  avec  $\rho_{XY} = 0.1$  et  $\rho_{XZ} = 0.3$  (Fig. 46). Sachant que le modèle Tau ne prend pas en compte l'influence de l'événement  $B$  sur l'événement  $C$ , nous avons aussi fait varier le coefficient de corrélation entre  $Y$  et  $Z$  :  $\rho_{YZ} = \{0.1, 0.4, 0.7, 0.9, 0.99\}$ . Comme le montre les figures de ces deux tests, la faible corrélation entre  $Y$  et  $Z$  permet un bon accord entre les deux modèles ; dès que la corrélation devient importante, le modèle gaussien s'éloigne du modèle Tau qui ne dépend pas de la corrélation  $\rho_{YZ}$ .



**Fig. 45** Exemple du non accord entre modèle gaussien et modèle Tau  $\tau = 1$  pour plusieurs valeurs de la corrélation entre  $Y$  et  $Z$  :  $\rho_{YZ} = \{0.1, 0.4, 0.7, 0.9, 0.99\}$ .



**Fig. 46** Exemple du non accord entre modèle gaussien et modèle Tau  $\tau = 1$  pour plusieurs valeurs de la corrélation entre  $Y$  et  $Z$  :  $\rho_{YZ} = \{0.1, 0.4, 0.7, 0.9\}$ .

Ces tests confirment également qu'une valeur fixe du paramètre  $\tau$  ne suffit pas pour reproduire le modèle gaussien. On peut alors se demander comment estimer les valeurs exactes du paramètre  $\tau$  pour reproduire le modèle gaussien. Pour ce faire, nous pouvons résoudre le problème inverse (Eq. 3.2). Nous avons effectué un test avec les paramètres  $y = -2.0$ ,  $z = 2.0$ ,  $\rho_{XY} = 0.1$ ,  $\rho_{XZ} = 0.3$ ,  $\rho_{YZ} = 0.7$ . D'après les résultats de ce test, nous pouvons constater que la reproduction exacte d'un modèle gaussien n'est possible qu'avec des valeurs de  $\tau$  estimées en chaque point de simulation (Fig. 47).

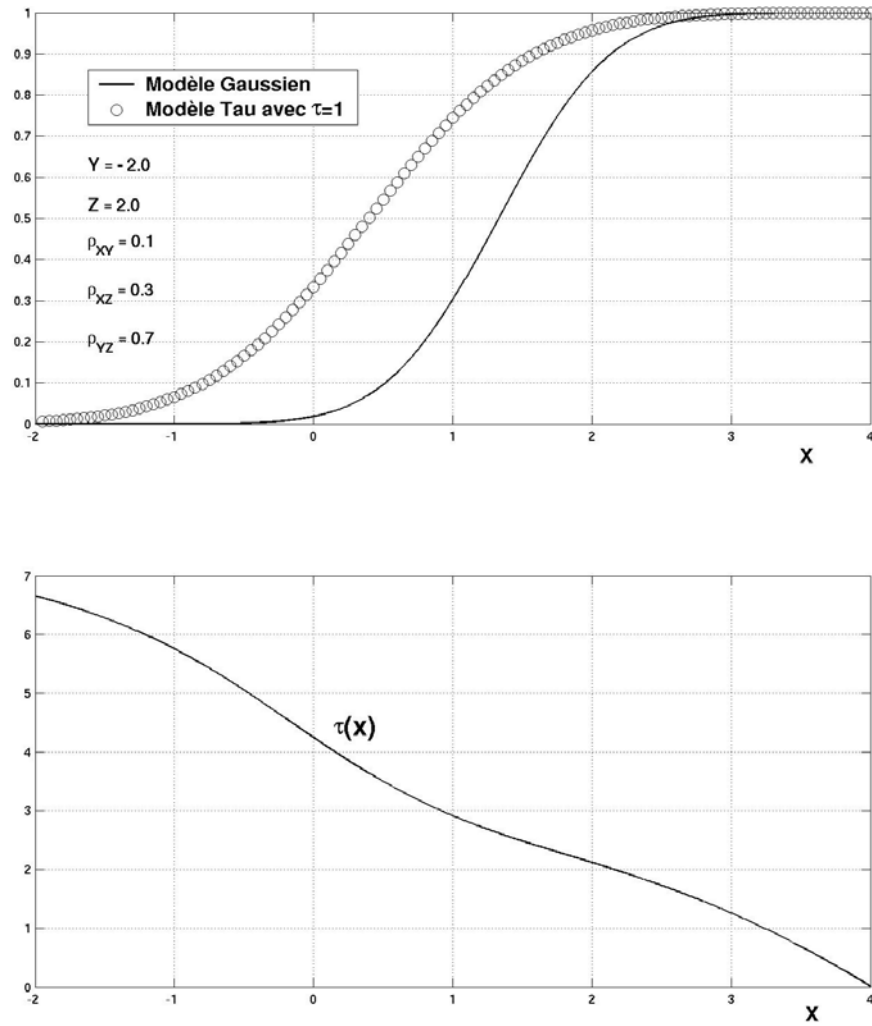


Fig. 47 Problème inverse : estimation du paramètre  $\tau$ , nécessaire pour reproduire le modèle gaussien.

**Modèle trigaussien seuillé.** Soient  $U, S_1, S_2$  et  $S_3$  quatre variables gaussiennes standards indépendantes. Soient aussi trois valeurs numériques  $\alpha, \beta$  et  $\gamma$  comprises entre 0 et 1. Nous définissons le vecteur gaussien  $(X, Y, Z)$  par :

$$\begin{cases} X = \sqrt{\alpha} \cdot U + \sqrt{1-\alpha} \cdot S_1 \\ Y = \sqrt{\beta} \cdot U + \sqrt{1-\beta} \cdot S_2 \\ Z = \sqrt{\gamma} \cdot U + \sqrt{1-\gamma} \cdot S_3 \end{cases} \quad 3.5$$

$(X, Y, Z)$  admet la matrice de covariance suivante :

$$\Lambda = \begin{pmatrix} 1 & \sqrt{\alpha\beta} & \sqrt{\alpha\gamma} \\ \sqrt{\alpha\beta} & 1 & \sqrt{\beta\gamma} \\ \sqrt{\alpha\gamma} & \sqrt{\beta\gamma} & 1 \end{pmatrix} \quad 3.6$$



Les événements qui nous intéressent sont maintenant  $A = \{X < x\}$ ,  $B = \{Y < y\}$  et  $C = \{Z < z\}$ . Calculons tout d'abord la probabilité conditionnelle :

$$P[A | B, C] = P[X < x | Y < y, Z < z] = \frac{P[X < x, Y < y, Z < z]}{P[Y < y, Z < z]}$$

Les probabilités conjointes s'évaluent par randomisation sur  $u$  en utilisant le fait que les variables  $U, S_1, S_2, S_3$  sont indépendantes. Notant  $g(\cdot)$  la densité gaussienne standard et  $G(\cdot)$  sa fonction de répartition, on a :

$$P[Y < y, Z < z] = \int_{-\infty}^{+\infty} g(u) \cdot G\left(\frac{y - \sqrt{\beta}u}{\sqrt{1 - \beta}}\right) \cdot G\left(\frac{z - \sqrt{\gamma}u}{\sqrt{1 - \gamma}}\right) du$$

$$P[X < x, Y < y, Z < z] = \int_{-\infty}^{+\infty} g(u) \cdot G\left(\frac{x - \sqrt{\alpha}u}{\sqrt{1 - \alpha}}\right) \cdot G\left(\frac{y - \sqrt{\beta}u}{\sqrt{1 - \beta}}\right) \cdot G\left(\frac{z - \sqrt{\gamma}u}{\sqrt{1 - \gamma}}\right) du$$

L'approximation par le modèle Tau requiert les calculs intermédiaires des probabilités conjointes suivantes :

$$P[A] = P[X < x]$$

$$P[A | B] = \frac{P[X < x, Y < y]}{P[Y < y]}$$

$$P[A | C] = \frac{P[X < x, Z < z]}{P[Z < z]}$$

où

$$P[X < x] = G\left(\sqrt{\alpha} \cdot U + \sqrt{1 - \alpha} \cdot S_1 < x\right) = \int_{-\infty}^{+\infty} g(u) \cdot G\left(\frac{x - \sqrt{\alpha}u}{\sqrt{1 - \alpha}}\right) du$$

$$P[X < x, Y < y] = \int_{-\infty}^{+\infty} g(u) \cdot G\left(\frac{x - \sqrt{\alpha}u}{\sqrt{1 - \alpha}}\right) \cdot G\left(\frac{y - \sqrt{\beta}u}{\sqrt{1 - \beta}}\right) du$$

$$P[X < x, Z < z] = \int_{-\infty}^{+\infty} g(u) \cdot G\left(\frac{x - \sqrt{\alpha}u}{\sqrt{1 - \alpha}}\right) \cdot G\left(\frac{z - \sqrt{\gamma}u}{\sqrt{1 - \gamma}}\right) du$$

Évaluons maintenant le degré d'approximation de la probabilité biconditionnelle par le modèle Tau avec  $\tau = 1$  donné par (Eq. 3.4). Un premier test est effectué pour les valeurs gaussiennes proches de sa moyenne  $y = 0.1$ ,  $z = 0.1$ . Nous avons fixé les corrélations  $\rho_{XY}$  et  $\rho_{XZ}$  à 0.5 tandis que nous avons utilisé deux valeurs pour  $\rho_{YZ}$  : 0.3 et 0.9. Dans les deux cas, nous observons un accord satisfaisant entre les deux modèles (Fig. 48). Un deuxième test a été réalisé avec les valeurs gaussiennes  $y = -2.0$ ,  $z = -2.0$  et avec le même jeu de coefficients de corrélation  $\rho_{XY} = 0.5$ ,  $\rho_{XZ} = 0.5$ ,  $\rho_{YZ} = \{0.3, 0.9\}$ . Pour les valeurs gaussiennes éloignées de leur moyenne, nous n'avons plus d'accord entre le modèle gaussien et son approximation par le modèle Tau (Fig. 49).

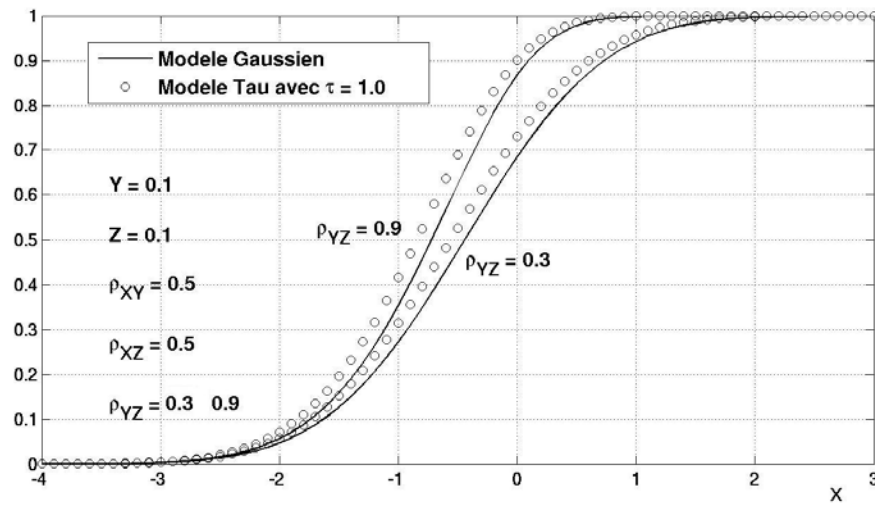


Fig. 48 Exemple de l'accord satisfaisant entre le modèle gaussien et le modèle Tau pour deux valeurs de la corrélation entre  $Y$  et  $Z$  :  $\rho_{YZ} = \{0.3, 0.9\}$ .

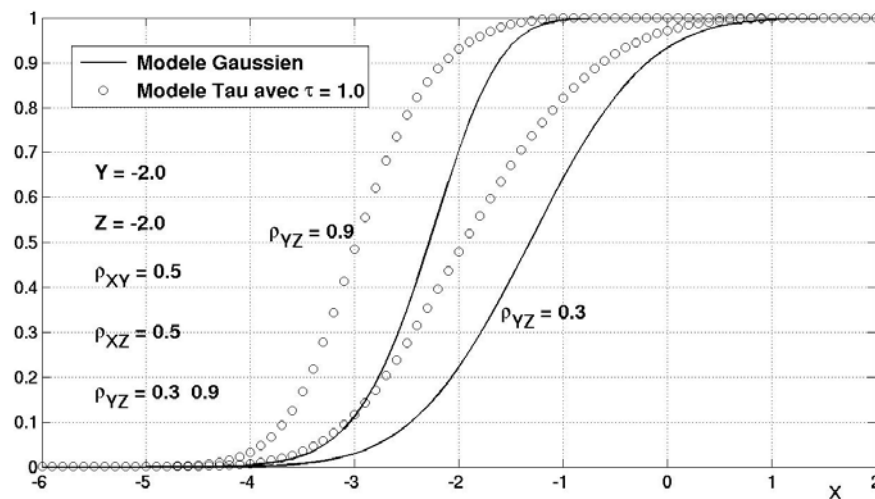


Fig. 49 Exemple du non accord entre le modèle gaussien et le modèle Tau pour deux valeurs de la corrélation entre  $Y$  et  $Z$  :  $\rho_{YZ} = \{0.3, 0.9\}$ .

Il est intéressant de remarquer que dans le cas du modèle gaussien classique trivariable l'approximation par le modèle Tau ne dépend pas de  $\rho_{YZ}$ , tandis que dans le cas du modèle gaussien seuillé trivariable cette même approximation suit le changement de  $\rho_{YZ}$  tout en maintenant un certain décalage. Cet effet est une spécificité du modèle gaussien seuillé : la dépendance linéaire est introduite non seulement entre les variables (Eq. 3.5) mais aussi entre les coefficients de corrélation (Eq. 3.6).

**Modèle autodual.** Un ensemble aléatoire est dit *autodual* lorsque sa loi spatiale coïncide avec celle de son complémentaire. Soient  $A$ ,  $B$  et  $C$  trois événements ponctuels. Ils vérifient donc :

$$P[A] = P[A^c], P[B] = P[B^c], P[C] = P[C^c] \quad 3.7$$

Il se trouve que l'autodualité permet d'exprimer la loi conjointe des 3 événements à celles des événements pris 2 à 2 :

$$\begin{aligned} P[A, B, C] &= P[A^c, B^c, C^c] \\ &= P[(A \cup B \cup C)^c] \\ &= 1 - P[A \cup B \cup C] \\ &= 1 - (P[A] + P[B] + P[C] - P[A, B] - P[A, C] - P[B, C] + P[A, B, C]) \end{aligned} \quad 3.8$$

Et compte tenue de 3.7, l'équation 3.8 devient :

$$P[A, B, C] = \frac{1}{2} \left[ P[A, B] + P[A, C] + P[B, C] - \frac{1}{2} \right] \quad 3.9$$

La probabilité biconditionnelle s'exprime donc :

$$\begin{aligned} P[A|B, C] &= \frac{1}{2} \frac{P[A, B] + P[A, C] + P[B, C] - \frac{1}{2}}{P[B, C]} \\ &= \frac{1}{2} \frac{P[A|B] + P[A|C] + P[B|C] - 1}{P[B|C]} \end{aligned} \quad 3.10$$

De son coté, l'approximation de la probabilité biconditionnelle par le modèle Tau s'écrit :

$$P^*[A|B, C] = \frac{1}{1 + \left(\frac{c}{a}\right)^\tau b} \quad 3.11$$

avec

$$a = \frac{1 - P[A]}{P[A]} \equiv 1, \quad b = \frac{1 - P[A|B]}{P[A|B]}, \quad c = \frac{1 - P[A|C]}{P[A|C]}$$

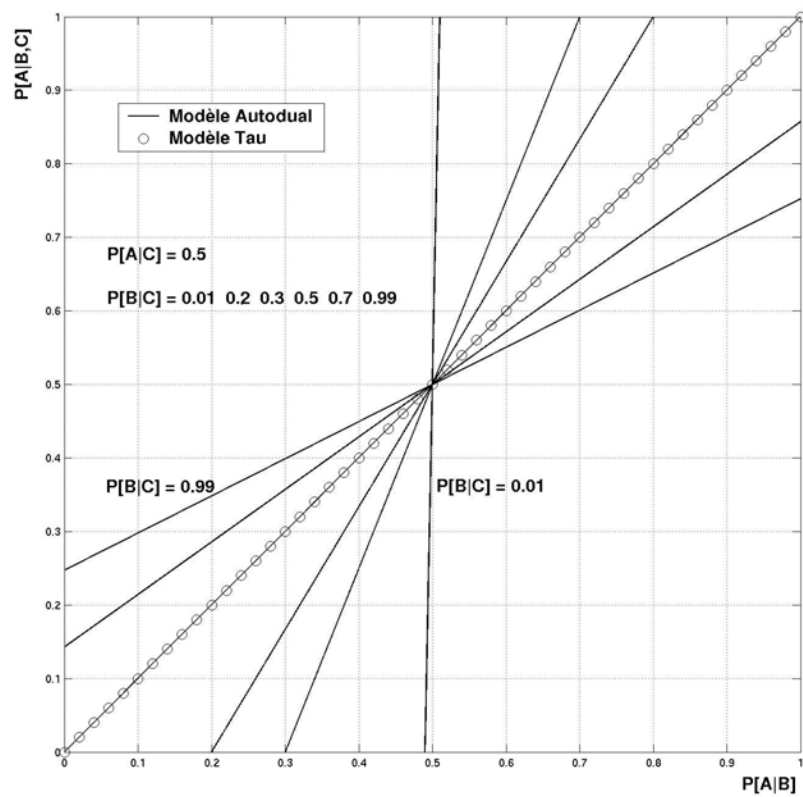
La probabilité (Eq. 3.10) est comparée à son approximation (Eq. 3.11) en faisant varier  $P[A|B]$  entre 0 et 1, et en fixant tous les autres paramètres. Le premier test est effectué avec les valeurs suivantes :  $P[A] = 0.5$ ,  $P[A|C] = 0.5$ ,  $P[B|C] = \{0.01 \ 0.2 \ 0.3 \ 0.5 \ 0.7 \ 0.99\}$ .

Il s'avère que les deux modèles ne coïncident que lorsque  $P[B|C] = 0.5$  (Fig. 50). Il est aussi intéressant de mentionner que, dans ce cas particulier, le modèle Tau n'est pas affecté par la valeur du paramètre  $\tau$ , car la proportion  $(a/c)^\tau = 1^\tau$  reste toujours égale à 1. C'est la raison pour laquelle il est impossible d'estimer le paramètre  $\tau$  (censé exprimer l'influence de l'événement  $B$  sur  $C$ ) en dépit d'une connaissance exhaustive de toutes les probabilités conditionnelles :

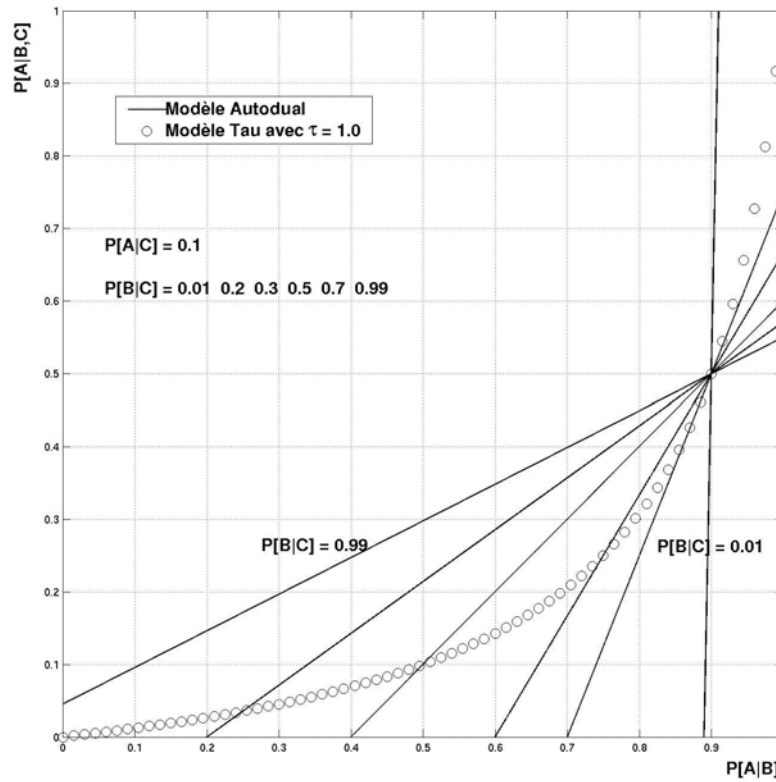
$$\tau = \frac{\log \frac{x}{b}}{\log \frac{a}{c}} = \frac{\log \frac{x}{b}}{\log 1}$$

Le cas  $P[A|C]=0.5=P[A]$  correspond à la situation où  $C$  n'apporte aucune information sur la connaissance de  $A$ . Ceci n'empêche d'ailleurs pas  $C$  d'apporter une contribution significative à l'information de  $B$  :  $P[B|C]=0.99$ .

Un deuxième test a été effectué, dans lequel toutes les probabilités ont été maintenues, hormis  $P[A|C]=0.1$ . Là, il n'a apparu aucun accord entre le modèle et son approximation (Fig. 51).



**Fig. 50 Exemple du non accord entre le modèle autodual et le modèle Tau.**  
Les paramètres utilisés sont  $P[A]=0.5$ ,  $P[A|C]=0.5$  et  $P[B|C]=\{0.01, 0.2, 0.3, 0.5, 0.7, 0.99\}$ .



**Fig. 51 Exemple du non accord entre le modèle autodual et le modèle Tau. Les paramètres utilisés sont  $P[A] = 0.5$ ,  $P[A|C] = 0.1$  et  $P[B|C] = \{0.01, 0.2, 0.3, 0.5, 0.7, 0.99\}$ .**

### 3.3.3 Application

Revenant aux notations habituelles de cette étude, la probabilité biconditionnelle est approximée par une fonction des probabilités nonconditionnelle et monoconditionnelles :

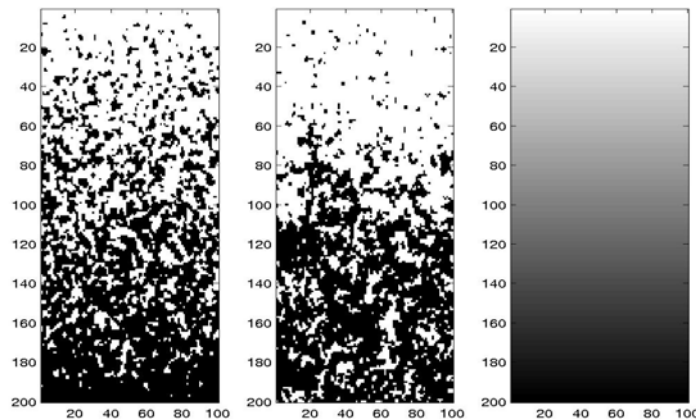
$$P[Z(u_0)|Z(\mathbf{u})=\mathbf{z}, S(u_0)=s] = f_\tau \left\{ P[Z(u_0)], P[Z(u_0)|Z(\mathbf{u})=\mathbf{z}], P[Z(u_0)|S(u_0)=s] \right\}$$

L'algorithme MP conditionné aux données auxiliaires à l'aide du modèle Tau consiste d'abord à évaluer la probabilité  $P[Z(u_0)|Z(\mathbf{u})=\mathbf{z}]$  puis à approximer la probabilité  $P[Z(u_0)|Z(\mathbf{u})=\mathbf{z}, S(u_0)=s]$  par la formule (3.1).

Voici maintenant trois exemples d'application bidimensionnels.

Le premier exemple est basé sur une **fonction gaussienne seuillée** (même exemple que pour la méthode de classement). La proportion locale de faciès varie le long de la verticale

sur l'image d'apprentissage. L'objectif de ce test est de reproduire cette non stationnarité (Fig. 52). En utilisant l'approximation du modèle Tau, nous n'avons besoin ni de classer les données auxiliaires, ni d'introduire l'image d'apprentissage auxiliaire. Disposant seulement de la carte de contrainte spatiale, nous conditionnons la réalisation MP par la proportion des faciès avec l'impact  $\tau = 1$ . Comme nous pouvons le constater, la répartition de faciès qui fournit la réalisation ne correspond pas à la carte de proportion. De plus, la proportion globale de faciès de la simulation (0.42) n'est pas celle de l'image d'apprentissage (0.5).



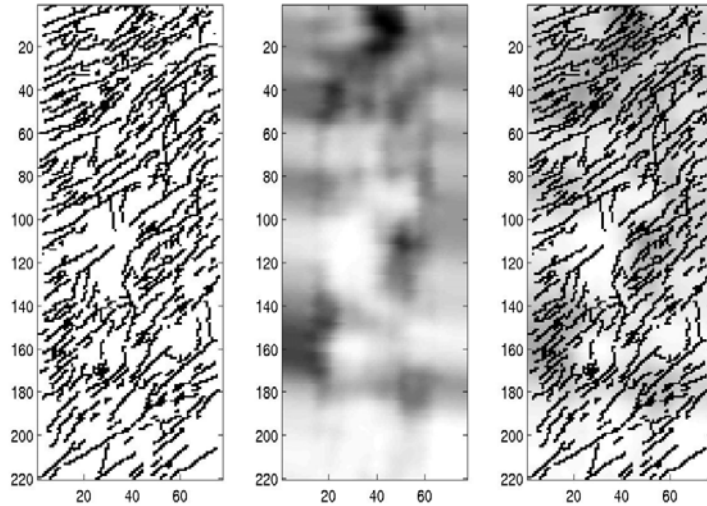
**Fig. 52 Simulation MP non stationnaire. Exemple sur une fonction gaussienne seuillée : l'image d'apprentissage (à gauche), la réalisation du modèle Tau avec  $\tau = 1$  (au centre) conditionnée par la carte de proportion locale de faciès (à droite). La proportion globale de faciès de l'image d'apprentissage est 0.5, celle de la réalisation est 0.42.**

**Réseau de fractures.** Les données initiales du deuxième exemple sont présentées en Fig. 53, notamment, le réseau de fractures à grande échelle (l'image d'apprentissage 77x220) et la carte de densité de fracturation (la contrainte). Cet exemple a déjà été vu dans le cadre stationnaire (§ 2.3.5). Ces mêmes données vont maintenant servir à simuler la répartition spatiale non stationnaire d'un réseau de fractures.

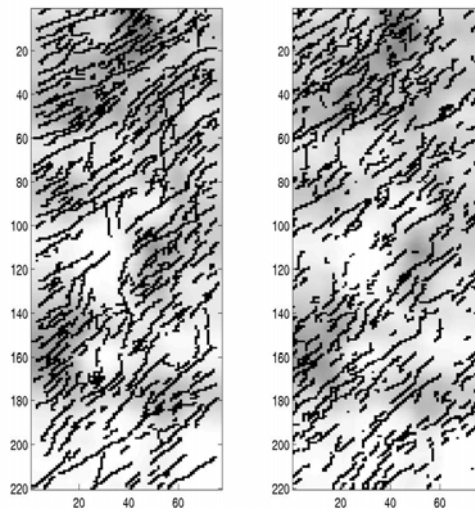
La Fig. 54 montre la simulation obtenue en prenant  $\tau = 1$ . La réalisation MP permet de distinguer des zones de densités locales différentes. Elles restituent assez bien la carte de densité locale. La densité globale de fracturation est de plus respectée. Cette réalisation peut donc être considérée comme acceptable pour notre test de simulation non stationnaire MP à l'aide du modèle Tau.

Il est tout à fait possible d'envisager des objets ponctuels assez nombreux dans les zones de faible densité. Ces objets représentent des fractures potentielles mais tronquées, puisque la densité locale ne favorise pas leur présence. C'est la raison pour laquelle nous avons cherché à réduire l'impact de la densité pour préserver la structure géométrique des

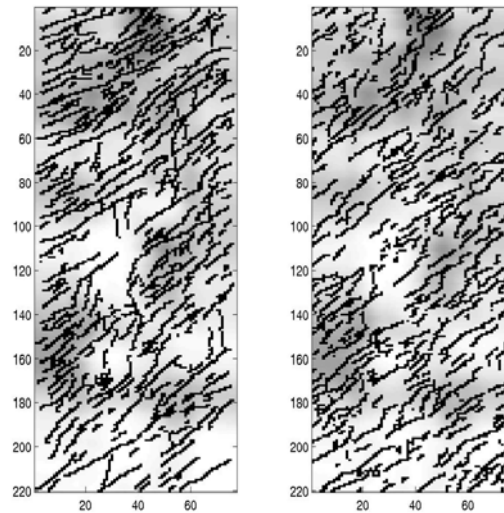
fractures. La Fig. 55 montre une simulation obtenue avec  $\tau = 0.5$ . Dans ce cas, nous pouvons voir que la structure géométrique des fractures est bien reproduite, tandis que la carte de densité n'est pas bien respectée.



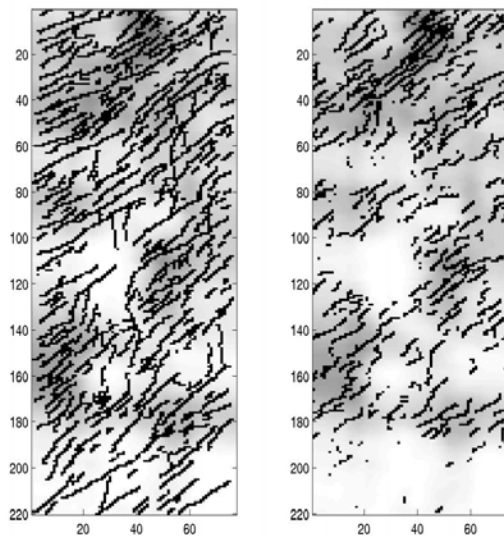
**Fig. 53** Image d'apprentissage à gauche, carte de densité au centre, superposition à droite.



**Fig. 54** Simulation MP non stationnaire. Exemple sur le réseau de fractures : l'image d'apprentissage à gauche, la réalisation du modèle Tau à droite, les deux sont superposées à la carte de densité. La proportion globale de faciès de l'image d'apprentissage égale à 0.28. La proportion globale de faciès de la réalisation est 0.28.  $\tau=1.0$ .

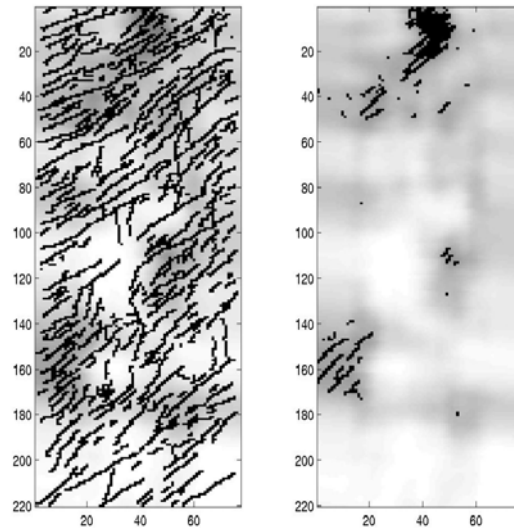


**Fig. 55 Simulation MP non stationnaire. Exemple sur le réseau de fractures :** l'image d'apprentissage à gauche, la réalisation du modèle Tau à droite, les deux sont superposées sur la carte de densité. La proportion globale de faciès de l'image d'apprentissage égale à 0.28. La proportion globale de faciès de la réalisation est 0.29.  $\tau=0.5$ .



**Fig. 56 Simulation MP non stationnaire. Exemple sur le réseau de fractures :** l'image d'apprentissage à gauche, la réalisation du modèle Tau à droite, les deux sont superposées sur la carte de densité. La proportion globale de faciès de l'image d'apprentissage égale à 0.28. La proportion globale de faciès de la réalisation est inférieure à 0.28.  $\tau=2.0$ .





**Fig. 57 Simulation MP non stationnaire. Exemple sur le réseau de fractures : l'image d'apprentissage à gauche, la réalisation du modèle Tau à droite, les deux sont superposées sur la carte de densité. La proportion globale de faciès de l'image d'apprentissage égale à 0.28. La proportion globale de faciès de la réalisation est inférieure à 0.28.  $\tau=10.0$ .**

Inversement, si on augmente l'impact des données auxiliaires ( $\tau = 2.0$  ou  $10.0$ ) pour mieux respecter la carte de densité, on obtient les objets géométriques (Fig. 56-57) qui ne figurent pas sur l'image d'apprentissage. En outre, la densité globale, qui dépend du paramètre  $\tau$  de manière non linéaire, est peu contrôlable.

**Segments orientés.** Dans les deux exemples que nous venons de voir, la fonction auxiliaire fournit une proportion locale de faciès ; l'exemple qui suit montre une intégration de données auxiliaires qui sont des orientations locales de segments.

L'image d'apprentissage principale, de taille 100x200, est une réalisation d'un schéma booléen de segments. L'orientation de ces derniers évolue graduellement de la verticale à l'horizontale en passant du bas au haut de l'image (Fig. 58A). Moyennant la convention « horizontal=0 » et « verticale =1 », cette orientation peut être utilisée aussi bien pour créer une image d'apprentissage auxiliaire que pour fournir un jeu de contraintes (Fig. 58B).

Pour pouvoir appliquer le Modèle Tau, il nous faut connaître la probabilité monoconditionnelle  $P[A | C] = P[Z(u_0) | S(u_0) = s]$ . Nous reproduisons en Fig. 58C la carte de probabilité  $P[Z(u_0) | S(u_0) = s]$  déduite des deux images d'apprentissages (Fig. 58A et B) qui, selon l'algorithme du modèle Tau, doit être utilisée pour intégrer l'orientation des segments. Cette carte contient de l'information au sujet de la proportion des faciès, mais aucune information sur l'orientation des segments.

Un résultat de simulation non stationnaire MP à l'aide du modèle Tau est présenté (Fig. 59). Cet exemple illustre l'impossibilité de ce modèle à reproduire l'orientation spatiale.

Cela tient à ce qu'il ne fait qu'intégrer l'information contenue dans la probabilité monoconditionnelle  $P[Z(u_0) | S(u_0) = s]$ , sans tenir compte de l'information de voisinage. Or l'orientation d'un segment ne se lit pas au niveau ponctuel. Cette situation est analogue à celle où les données auxiliaires n'apportent pas d'information à la présence de faciès  $P[A | C] = P[A]$  mais seulement à sa présence dans un voisinage  $P[A | B, C] \neq P[A | B]$  (§ 3.3.2 modèle autodual :  $P[A | C] = 0.5$ ).

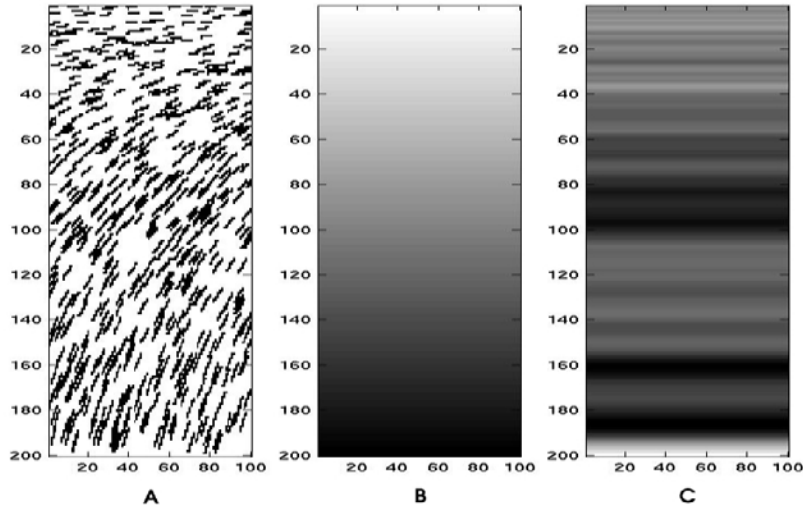


Fig. 58 (A) Image d'apprentissage principale; (B) image d'apprentissage auxiliaire; (C) carte (déduite des deux premières) représentant la probabilité monoconditionnelle  $P[Z(u_0) | S(u_0) = s]$ .

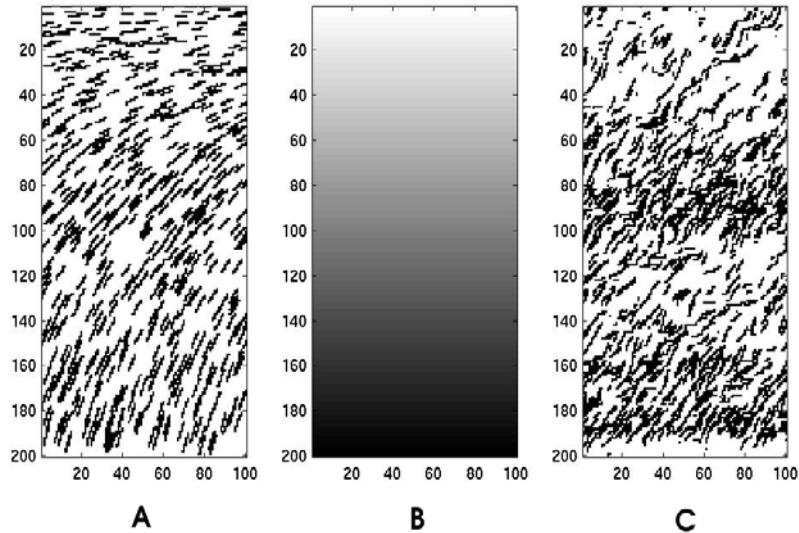


Fig. 59 (A) Image d'apprentissage principale ; (B) image d'apprentissage auxiliaire et la contrainte spatiale; (C) réalisation du modèle Tau avec  $\tau=1$ .

### 3.3.4 Remarques

Le test du modèle Tau sur trois modèles théoriques montre qu'une valeur constante du paramètre  $\tau$  ne suffit pas à approximer correctement les relations complexes entre les différents événements. Par ailleurs, pour estimer  $\tau$  en chaque point de simulation, il faut connaître la probabilité  $P[A | B, C]$  elle-même (Eq. 3.2).

Le modèle Tau suppose connue la probabilité conditionnelle  $P[A | C] = P[Z(u_0) | S(u_0) = s]$ , sans faire explicitement appel à une image d'apprentissage auxiliaire. Ceci peut être considéré comme un atout, car dans certains cas, l'information conditionnelle  $P[A | C]$  est directement fournie par la sismique. Dans les autres cas, cette information devra être extraite d'une image d'apprentissage auxiliaire.

Les exemples que nous avons présentés montrent que la simulation MP à l'aide du modèle Tau ne fournit pas de résultats satisfaisants. Ces derniers dépendent trop du choix du paramètre  $\tau$ . En conditionnant avec des données auxiliaires spatiales, le modèle Tau ne prend en compte que la proportion ponctuelle de faciès, ce qui ne lui permet pas de restituer des caractéristiques géométriques régionales (e.g. orientation).

Dans deux des exemples proposés (le modèle autodual avec  $P[A | C] = 0.5$  et segments orientés), le modèle Tau s'avère impuissant à délivrer des simulations MP correctes. Ces deux exemples peuvent être extrapolés à tous les cas où l'événement  $C$  apporte peu d'information à  $A$  mais plus à  $B$  :  $P[A | C] \approx P[A]$ ,  $P[B | C] \neq P[B]$ .

Pour résumer, les inconvénients principaux de cette approche sont : sa portée limitée ; son recours à un modèle probabiliste analytique ; son contrôle insuffisant de la densité globale ; son impossibilité de respecter à la fois la géométrie des objets et la contrainte auxiliaire ; sa dépendance envers le type de contraintes spatiales.

## 4 Simulation non stationnaire : l'approche MPNST

Les inconvénients que nous venons de mentionner nous amènent à développer une nouvelle approche de simulation non stationnaire qui possède les caractéristiques suivantes :

- 1) indépendance d'un modèle probabiliste analytique ;
- 2) possibilité de conditionnement par des données auxiliaires représentant la non stationnarité ;
- 3) pas de réduction ni de transformation de données auxiliaires (continues) ;
- 4) indépendance du type de données auxiliaires.

Dans ce chapitre nous présentons l'approche que nous avons développée et qui nous permet de reproduire des structures géométriques complexes sous contrainte spatiale. Nous commençons par expliciter le principe de notre approche et l'algorithme de base qui met en jeu la notion de support d'une contrainte spatiale. Nous détaillons ensuite la mise en oeuvre de notre approche qui s'appuie de nouveau sur l'étape de scanning et l'étape de simulation à l'aide de l'arbre de stockage. Nous présentons finalement quelques exemples de validation (fonction gaussienne seuillée, réseaux de fractures, agrégat de segments orientés) tout en comparant ces résultats avec les deux approches existantes.

Cette nouvelle approche étant destinée à effectuer des simulations MP dans un cadre non (N) stationnaire (ST), nous proposons de lui attribuer l'acronyme MPNST.

### 4.1 Présentation Générale

#### 4.1.1 Principe

Tout comme pour la méthode de classement, l'approche MPNST consiste à utiliser, en plus d'une image d'apprentissage principale, une image d'apprentissage auxiliaire (Fig. 60). Nous avons besoin de cette dernière pour établir une liaison entre la géométrie de la fonction principale  $Z$  et les valeurs de la fonction auxiliaire  $S$  correspondante. L'approche MPNST repose ainsi sur trois idées :

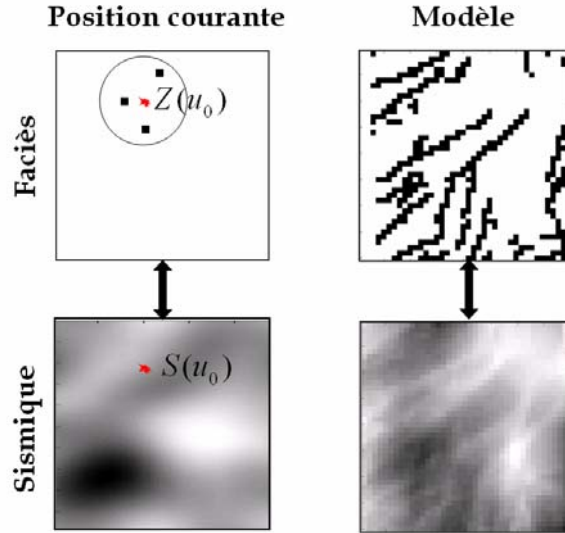
- 1) extraire la fréquence d'apparition de structures géométriques avec prise en compte de valeurs auxiliaires à partir de deux images d'apprentissage;

- 2) appliquer un critère de compatibilité entre la géométrie et la contrainte spatiale (valeurs auxiliaires);
- 3) utiliser les statistiques MP de géométrie compatible avec la contrainte spatiale pour effectuer une simulation non stationnaire.

Il existe en effet un lien entre la fonction principale et la fonction auxiliaire de sorte que la valeur auxiliaire  $S(u_0)$  « reflète » les valeurs prises par  $Z$  sur ensemble des points voisins autour de  $u_0$ . Appelons cet ensemble de points  $O(u_0)$  un *support*. Par exemple, dans le cas où la fonction auxiliaire représente la proportion locale de 2 faciès ( $Z \in \{0,1\}$ ), le support peut être défini comme le domaine sur lequel s'effectue la moyenne locale :

$$S(u_0) = \sum_{u \in O(u_0)} Z(u) \quad \text{Eq 4.1}$$

Le support dépend de la géométrie de la fonction principale et de la résolution de la fonction auxiliaire. Il est défini au cas par cas.



**Fig. 60** Espace de simulation avec sa carte de contrainte spatiale à gauche ; deux images d'apprentissage à droite.

Ainsi, nous pouvons établir une relation entre un état de la fonction principale  $Z$  sur le support  $O(u_0)$  et une valeur auxiliaire :  $Z\{O(u_0)\} \xrightarrow{F} S(u_0)$ . Cette relation dépend de la nature des données auxiliaires (la proportion, l'orientation ...) et aussi du procédé par lequel la valeur auxiliaire est déduite. Pour vérifier la compatibilité d'un état de support avec une contrainte spatiale, il nous faut à chaque fois calculer la valeur auxiliaire du support selon une méthode préalablement choisie  $F(Z\{O(u_0)\}) = ?$  et la comparer, ensuite, avec la contrainte spatiale  $S(u_0)$ . L'utilisation d'une telle relation dans l'algorithme de l'approche MPNST nous oriente vers une problématique purement déterministe. Nous souhaiterions toutefois garder notre approche indépendante de la nature des données auxiliaires et établir plutôt une relation statistique.

En pratique, le même état de support  $Z\{O(u_0)\}$  se retrouve plusieurs fois dans l'image d'apprentissage avec des valeurs auxiliaires  $\{S_1, S_2, S_3, \dots\}$ . La relation entre un état de support et ses valeurs auxiliaires peut donc être établie par la loi conditionnelle  $P[S < s | Z\{O(u_0)\}]$ . Cette loi extraite de deux images d'apprentissage (principale et auxiliaire) nous permet d'effectuer une simulation MP non stationnaire comme nous allons le voir ci-dessous.

### 4.1.2 Algorithme

L'algorithme MPNST est séquentiel, tout comme l'algorithme MP. Avant de commencer, la taille et la configuration du support doivent être choisis en fonction de la géométrie de  $Z$  et de la résolution de  $S$ .

Soit  $u_0$  le point courant à simuler, et sois  $\mathbf{u}$  le voisinage de ses points déjà simulés. Sur ce voisinage on a  $Z(\mathbf{u}) = \mathbf{z}$ . Soient  $w_1, w_2, w_3, \dots$  les positions des répliques de  $Z(\mathbf{u}) = \mathbf{z}$  dans l'image d'apprentissage. Les valeurs de  $Z$  sur les supports  $O(w_1), O(w_2), O(w_3), \dots$  coïncident sur les translatés de  $\mathbf{u}$  mais pourront être différentes aux autres points. Désignons aussi par  $S_1, S_2, S_3, \dots$  les valeurs auxiliaires attachées aux points  $w_1, w_2, w_3, \dots$ . Nous commençons par construire la loi conditionnelle  $P[S < s | Z\{O(w)\}]$  associée à chaque valeur de support. Puis nous appliquons à cette loi un *critère de compatibilité* (détaillé dans le paragraphe ci-dessous), ce qui nous permet d'identifier les supports compatibles à la contrainte spatiale  $S(u_0) = s$ . Les supports sont sélectionnés pour finalement fournir une loi MP de la forme  $P[Z(u_0) | Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s]$ , c'est-à-dire conditionnée par les valeurs prise au voisinage de  $u_0$  et la contrainte spatiale  $S(u_0) = s$ . Une valeur  $z_0$  est attribuée au point  $u_0$  selon cette loi conditionnelle. On passe alors au point suivant. La simulation s'arrête quand tous les points ont été traités. Voici l'algorithme qui correspond à cette approche :

- (i) Choisir le point non encore simulé  $u_0 \in E$ , extraire l'état de son voisinage  $Z(\mathbf{u}) = \mathbf{z}$  et sa contrainte spatiale  $S(u_0) = s$  ;
- (ii) Enregistrer tous les supports qui contiennent les répliques  $Z(\mathbf{u}) = \mathbf{z}$  :

$$W = \{w \in I : \mathbf{z} \subseteq Z\{O(w)\}, \mathbf{u} \subseteq O(w)\}$$

- (iii) Pour chaque état de support  $Z\{O(w)\}$  enregistré dans  $W$ , construire la loi  $P[S < s | Z\{O(w)\}]$  selon les valeurs auxiliaires  $\{S_1, S_2, S_3, \dots\}$  associées à cet état. Sélectionner ensuite les répliques selon un certain critère

de compatibilité (défini plus loin) :

$$\tilde{W} = \{w \in W : Z\{O(w)\} \xleftrightarrow{\text{critère}} S(u_0)\}$$

(iv) Si  $\tilde{W} = \emptyset$ , faire  $\mathbf{u} = \mathbf{u} \setminus u_n$  et retourner en (ii)

(v) Simuler la valeur de  $Z$  au point  $u_0$  selon la loi conditionnelle :

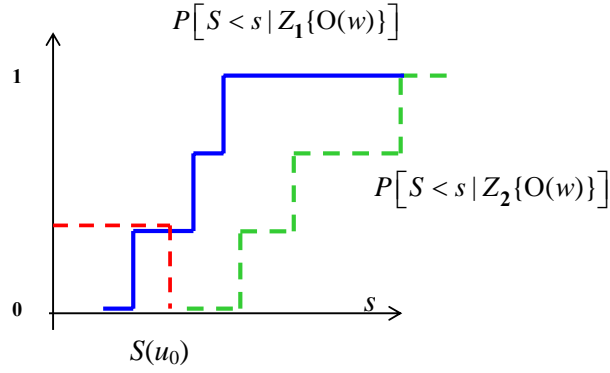
$$P[Z(u_0) = k \mid Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s] = \frac{\#\{w_0 \in \tilde{W} : Z(w_0) = k\}}{\#\{w_0 \in \tilde{W}\}}, \quad k = 1, 2, \dots, K$$

(vi) Retourner en (i)

### 4.1.3 Critère de compatibilité

Le critère de compatibilité est un indicateur statistique qui nous permet d'identifier des états de support compatibles à la contrainte spatiale. Par exemple, dans le cas où la fonction auxiliaire est une orientation régionale de segments, le critère nous indique le degré de compatibilité entre tel état de support et telle orientation. Comme il n'y a pas une manière universelle de définir ce critère, nous en proposons plusieurs.

Rappelons, que chaque état de support  $Z\{O(w)\}$  trouvé dans l'image d'apprentissage est associé à une loi de valeurs auxiliaires :  $P[S < s \mid Z\{O(w)\}]$ . Cette loi obtenue empiriquement permet de répondre à la question : « quelle est la probabilité d'avoir telle valeur auxiliaire (telle caractéristique) pour tel état de support (telle géométrie) ». A cette fin, pour simuler la fonction principale au point  $u_0$ , seuls les états de support de probabilité non nulle seront utilisés, ceux de probabilité nulle ne sont pas considérés. L'exemple illustré (Fig. 61) nous montre deux lois de fonction auxiliaire conditionnée par deux états de support. Selon le principe d'élimination des états non compatibles à la contrainte spatiale  $S(u_0) = s$ , nous éliminons l'état  $Z_1\{O(w)\}$  et gardons l'état  $Z_2\{O(w)\}$  pour la simulation ultérieure. S'appuyant sur ce principe, nous pouvons développer des critères plus au moins stricts conformément à nos besoins.



**Fig. 61** Exemple des lois conditionnelles de la fonction auxiliaire  $S$  pour deux états de support différents  $Z_1\{O(w)\}$  et  $Z_2\{O(w_0)\}$ .

- 1) **Le critère d'intervalle.** On dit qu'un état de support  $Z\{O(w)\}$  est compatible avec la contrainte spatiale  $S(u_0) = s$  si la probabilité  $P[S > s | Z\{O(w)\}]$  est strictement entre 0 et 1. Autrement dit, nous acceptons l'état du support, si la valeur  $S(u_0)$  est supérieure au quantile minimal  $q_{\min} = P^{-1}(0)$  et inférieure au quantile maximal  $q_{\max} = P^{-1}(1)$  de la loi :  $q_{\min} < s < q_{\max}$ .
- 2) **Le critère renforcé.** Soit  $\varepsilon$  une valeur positive préalablement choisie. L'état du support est considéré compatible avec la contrainte spatiale  $S(u_0) = s$  si  $q_{\min} + \varepsilon < s < q_{\max} - \varepsilon$ .
- 3) **Le critère relaxé.** Soit  $\varepsilon$  une valeur positive préalablement choisie. On dit qu'un état de support est compatible avec la contrainte spatiale  $S(u_0) = s$  si  $q_{\min} - \varepsilon < s < q_{\max} + \varepsilon$ .

Les états de support compatibles à la contrainte spatiale sont utilisés pour construire une loi probabiliste conditionnée à la fois au voisinage et à la contrainte. Mais parmi les supports choisis, certains peuvent être plus cohérents à la contrainte spatiale (la probabilité est plus importante) que d'autres. Il est alors logique de leur donner plus d'importance. Pour raffiner le conditionnement aux contraintes spatiales, il est possible de quantifier cette importance en attribuant un poids à chaque état de support. La loi conditionnelle est alors construite en utilisant ces poids. Nous présentons en l'Annexe C plusieurs procédés de construction de cette loi. Dans la suite de cette étude (sauf spécification contraire), nous exploitons le critère relaxé dans l'approche MPNST.



## 4.2 Mise en oeuvre

Pour la mise en oeuvre de l'approche MPNST, nous avons choisi de reprendre la structure arborescente (l'arbre de stockage) de Strebelle (2000) pour sa rapidité et sa facilité d'accès aux données statistiques. L'algorithme de notre approche traduit alors l'étape de scanning et l'étape de simulation.

Une implémentation efficace de notre approche nécessite de choisir correctement la configuration et la taille de support des données auxiliaires : une valeur auxiliaire dépend des valeurs prises aux points du support et uniquement en ces points. Pour obtenir des fréquences d'apparition de telle géométrie avec telle valeur auxiliaire, nous effectuons un scanning des images d'apprentissage à l'aide du support approprié. Un point clé de cette approche est d'avoir **la fenêtre d'exploration identique au support de fonction auxiliaire**.

Une fois l'image d'apprentissage scannée, la simulation peut débuter. Tout comme pour l'approche MP, l'espace de simulation est parcouru selon le chemin préalablement défini. Chaque point est simulé selon son voisinage (tout au moins sa partie visible dans le support) et sa contrainte spatiale. La simulation s'arrête quand tous les points ont été simulés.

Voici le diagramme d'algorithme MPNST auquel nous sommes parvenus (Fig. 62).

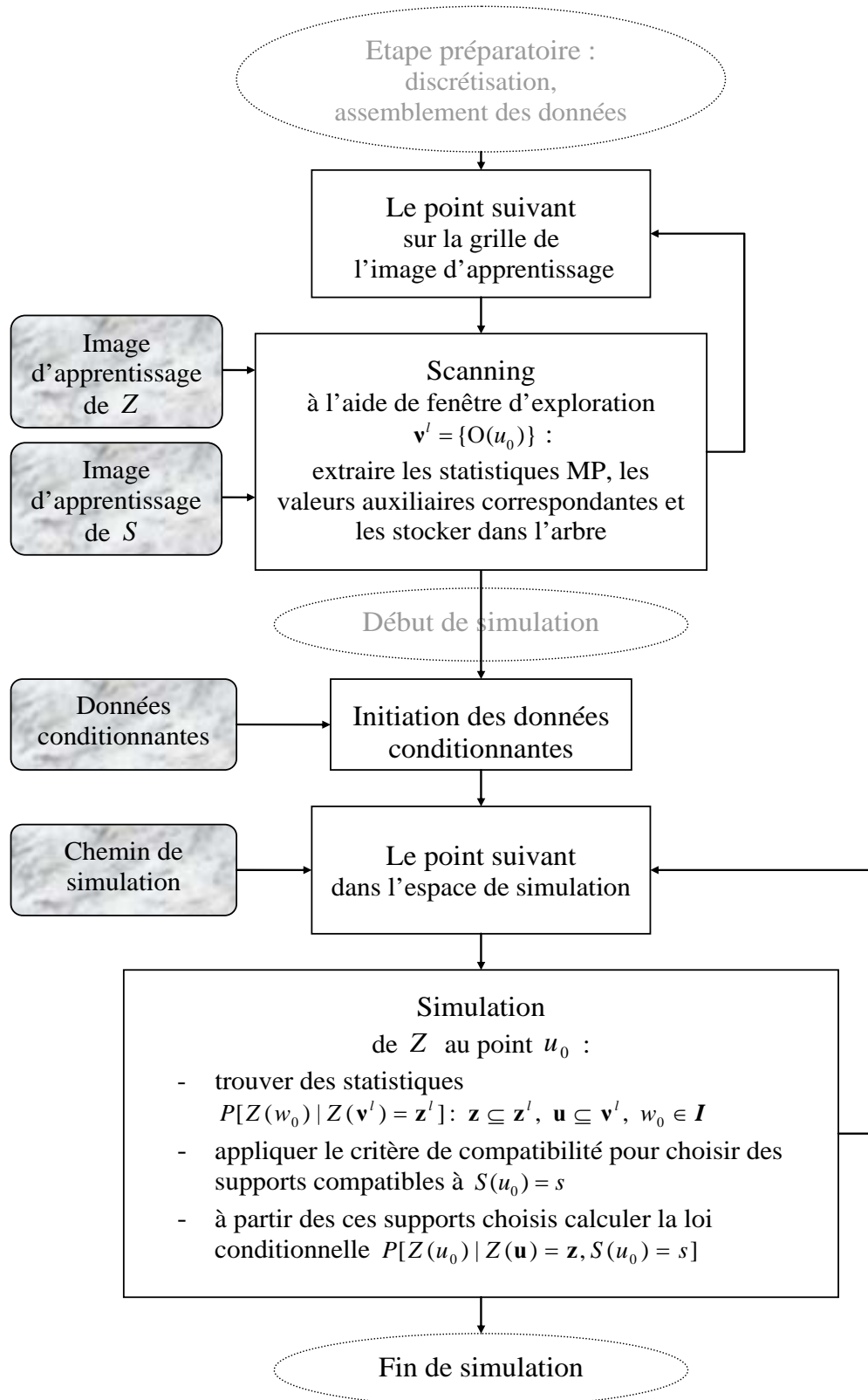


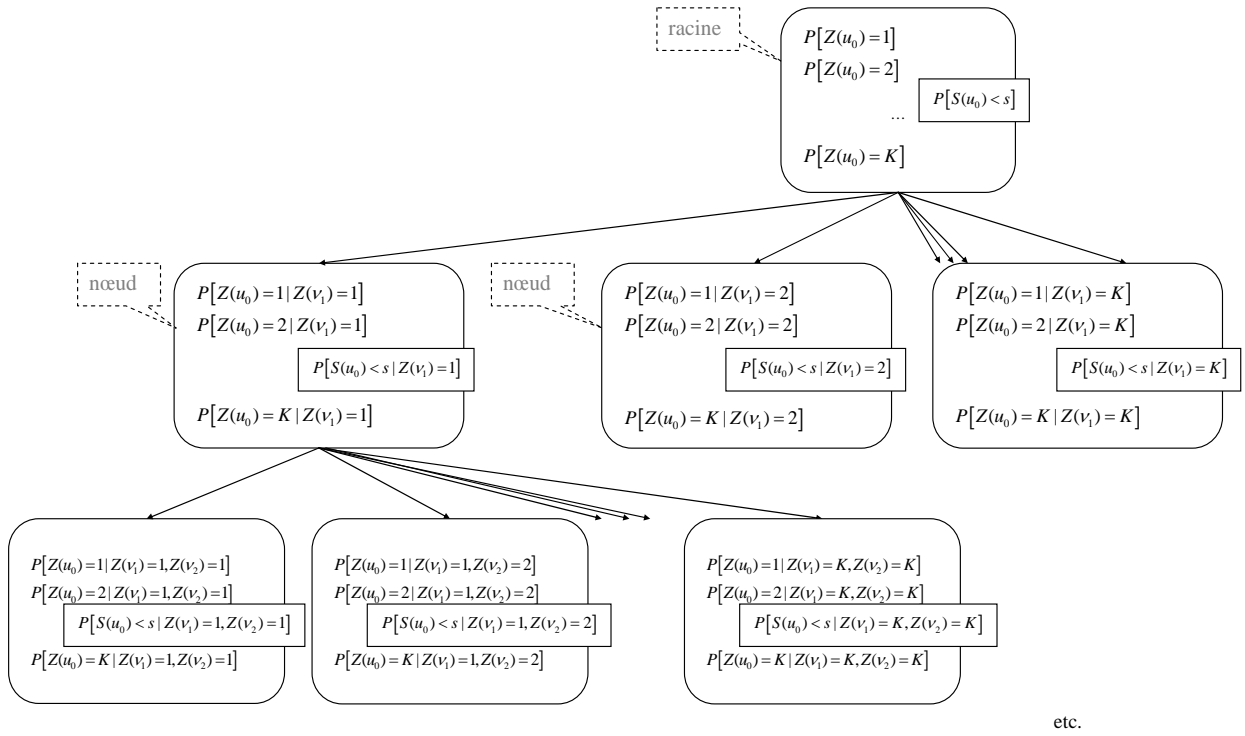
Fig. 62 Algorithme de simulation MPNST.

### 4.2.1 Scanning de l'image d'apprentissage

La première étape de l'approche MPNST consiste à calculer et à stocker les statistiques MP et les valeurs auxiliaires correspondantes. Cela s'effectue par le scanning de l'image d'apprentissage à l'aide d'une fenêtre d'exploration identique au support  $\mathbf{v}^L \equiv \mathbf{O}(u_0)$ .

A chaque noeud de l'arbre de stockage sont associés : 1) un état particulier  $Z(\mathbf{v}^l) = \mathbf{z}^l$  ; 2) les nombres de répliques de cet état pour les différentes valeurs du point central ; 3) les valeurs auxiliaires correspondant à cet état  $\{S_1, S_2, S_3 \dots\}$ . L'ensemble de cette information fournit deux lois empiriques :  $P[Z(u_0) | Z(\mathbf{v}^l) = \mathbf{z}^l]$ ,  $u_0 \in \mathbf{I}$  et  $P[S < s | Z(\mathbf{v}^l) = \mathbf{z}^l]$ .

Le schéma de l'arbre de stockage est présenté dans le cas général de  $K$  faciès (Fig. 63).



**Fig. 63 L'arbre de stockage pour les statistiques MP et les valeurs auxiliaires extraites des images d'apprentissage. Cas général de  $K$  faciès.**

### 4.2.2 Simulation

La simulation commence par le positionnement des données fixes et certaines dans l'espace de simulation. On parcourt ensuite le chemin de simulation (aléatoire ou non).

À chaque point de simulation  $u_0 \in E$  correspond la loi biconditionnelle suivante :

$$P[Z(u_0) | Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s] \quad \text{Eq 4.2}$$

Pour estimer la probabilité (Eq 4.2), on recourt à l'algorithme suivant :

- (i) Choisir le point  $u_0 \in E$  selon le chemin prédéfini, extraire l'état de son voisinage  $Z(\mathbf{u}) = \mathbf{z}$  et sa contrainte spatiale  $S(u_0) = s$  ;
- (ii) Choisir le niveau d'arborescence qui correspond à la taille minimum de la sous fenêtre contenant le voisinage simulé :

$$l = \min_{0 \leq l \leq L} \{\mathbf{u} \subseteq \mathbf{v}^l\}.$$

- (iii) Retenir dans l'ensemble des noeuds de ce niveau ceux qui coïncident avec l'état du voisinage  $Z(\mathbf{u}) = \mathbf{z}$  :

$$W = \{w_0 \in I : Z(\mathbf{v}^l) = \mathbf{z}^l, \mathbf{z} \subseteq \mathbf{z}^l\}.$$

- (iv) Pour chaque état de nœud  $\mathbf{v}^l$  enregistré dans  $W$ , construire la loi  $P[S < s | Z(\mathbf{v}^l) = \mathbf{z}^l]$ . Ensuite, trier les noeuds selon le critère de compatibilité retenu :

$$\tilde{W} = \{w_0 \in W : Z(\mathbf{v}^l) \xleftrightarrow{\text{critère}} S(u_0)\}$$

- (v) Si  $\tilde{W} = \emptyset$ , faire  $\mathbf{u} = \mathbf{u} \setminus \nu_l$  et retourner en (ii) ;

- (vi) Simuler la valeur de  $Z$  au point  $u_0$  selon la loi conditionnelle :

$$P[Z(u_0) = k | Z(\mathbf{u}) = \mathbf{z}, S(u_0) = s] = \frac{\#\{w_0 \in \tilde{W} : Z(w_0) = k\}}{\#\{w_0 \in \tilde{W}\}}, \quad k = 1, 2, \dots, K$$

- (vii) Retourner en (i)

## 4.3 Exemples de validation

Nous présentons ici la partie du travail qui consiste à valider le fonctionnement de notre approche et à la comparer avec deux autres approches de simulation MP non stationnaire : l'approche MP conditionnée à l'aide du modèle Tau et l'approche MP conditionnée à l'aide du classement des valeurs auxiliaires. Comme l'étude du modèle Tau l'a montré (3.3.3), les meilleurs résultats sont obtenus en prenant le paramètre  $\tau$  égal à 1. Nous utilisons cette valeur pour tous les tests numériques présentés ci-dessous. Nous proposons trois types de modèles qui joueront le rôle d'image d'apprentissage : une fonction gaussienne seuillée, un réseau de fractures et un schéma booléen de segments orientés.

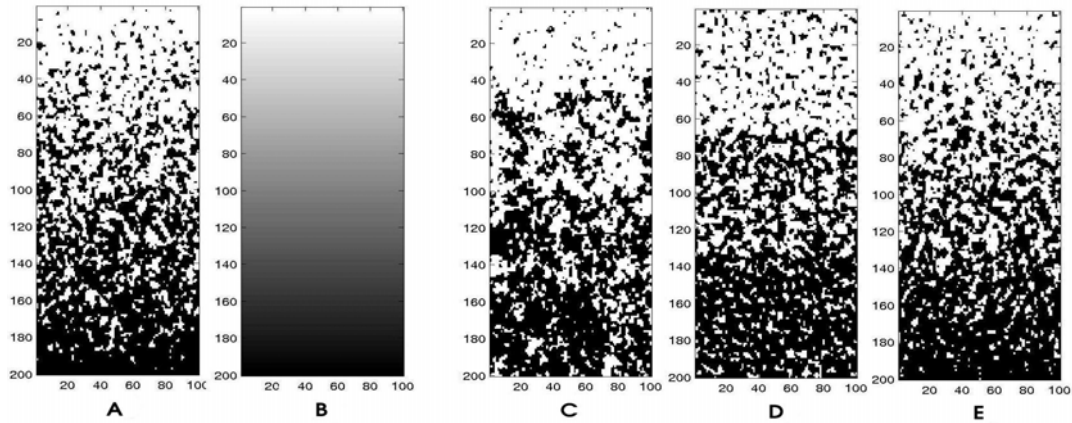
### 4.3.1 Fonction gaussienne seuillée

Nous prenons le même exemple de fonction gaussienne déjà étudié en § 3.2.2. Rappelons ses caractéristiques : l'image d'apprentissage principale est une réalisation d'une fonction gaussienne 2D de variogramme gaussien (portée 3). La fonction a été seuillée en 2 dégradés de telle façon que la proportion locale de faciès varie continûment de 1 à 0 le long de la verticale. Toutes les valeurs de densité locale apparaissent ainsi sur l'image d'apprentissage. Dans notre cas d'étude, cette proportion fournit à la fois l'image d'apprentissage auxiliaire et la contrainte spatiale.

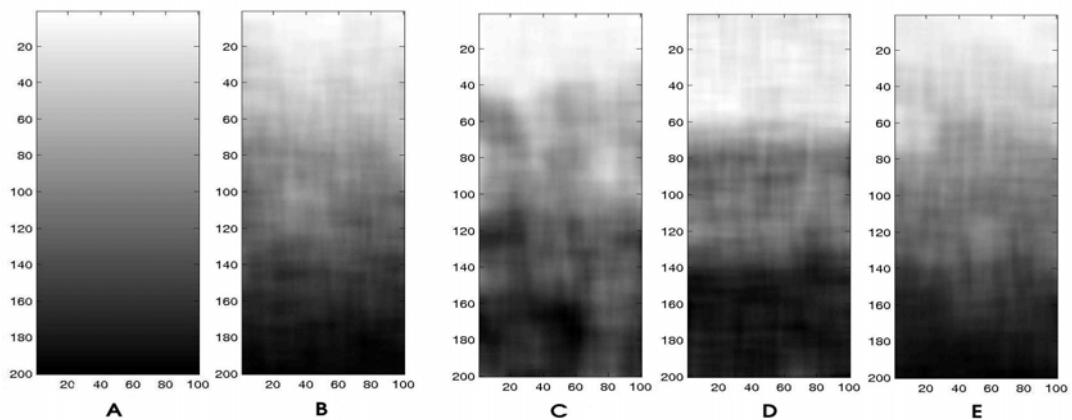
La comparaison des approches (Fig. 64) montre une différence significative entre la réalisation du modèle Tau et l'image d'apprentissage. La méthode de classement restitue mieux la géométrie mais l'effet de classement entraîne une compartimentation de l'espace de simulation en 3 zones (Fig. 64D). Pour s'assurer que la réalisation de l'approche MPNST est plus conforme à la contrainte spatiale que les deux approches existantes, nous avons calculé des cartes de moyenne locale (Fig. 65). Un autre test est effectué avec les mêmes images d'apprentissage mais avec une contrainte spatiale distincte (Fig. 66) dont la gamme des valeurs reste entre 0 et 1. Cette fois encore, la moyenne locale des réalisations a été comparée à la contrainte spatiale (Fig. 67).

Dans ces deux tests, nous pouvons remarquer que les trois approches produisent toutes des structures géométriques non stationnaires plus au moins conformes à la contrainte de proportion. Le modèle Tau fournit la réalisation de moins bonne qualité : la géométrie reproduite ne correspond pas à la contrainte spatiale. De ce point de vue, la méthode de classement permet de mieux reproduire la géométrie, mais présente néanmoins des

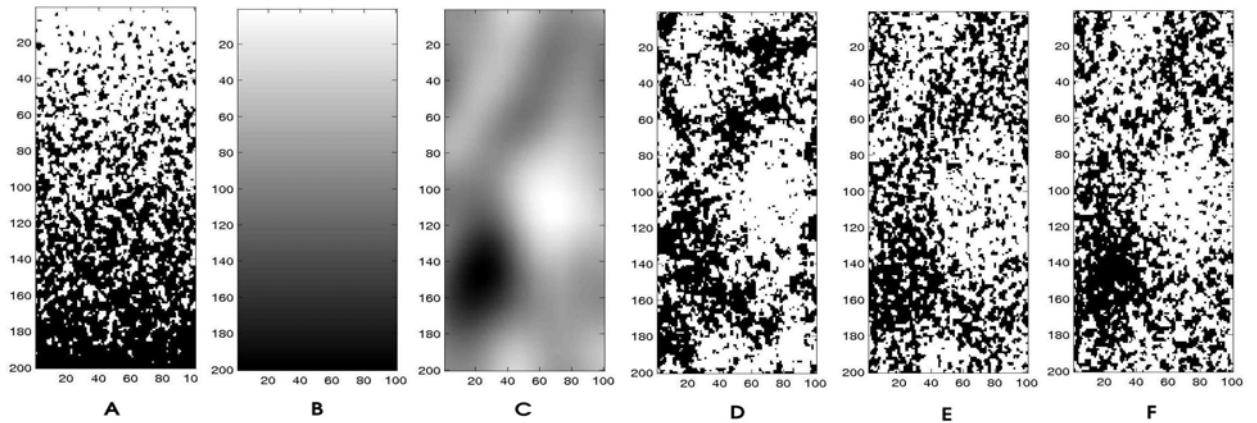
artefacts : la géométrie est homogène dans chaque zone de classement. L'approche MPNST permet une très bonne reproduction de la géométrie en conformité avec la carte de contrainte.



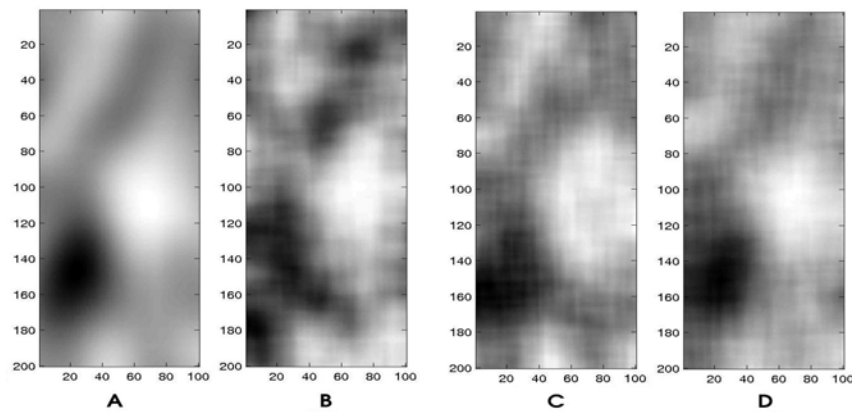
**Fig. 64** (A) Image d'apprentissage principale ; (B) image d'apprentissage auxiliaire et contrainte spatiale; (C) réalisation du modèle Tau ; (D) réalisation de la méthode de classement ; (E) réalisation de l'approche MPNST.



**Fig. 65** (A) Contrainte spatiale; moyenne mobile de : (B) l'image d'apprentissage principale; (C) la réalisation du modèle Tau ; (D) la réalisation de la méthode de classement ; (E) la réalisation de l'approche MPNST.



**Fig. 66** (A) Image d'apprentissage principale ; (B) image d'apprentissage auxiliaire ; (C) contrainte spatiale ; (D) réalisation du modèle Tau ; (E) réalisation de la méthode de classement ; (F) réalisation de l'approche MPNST.



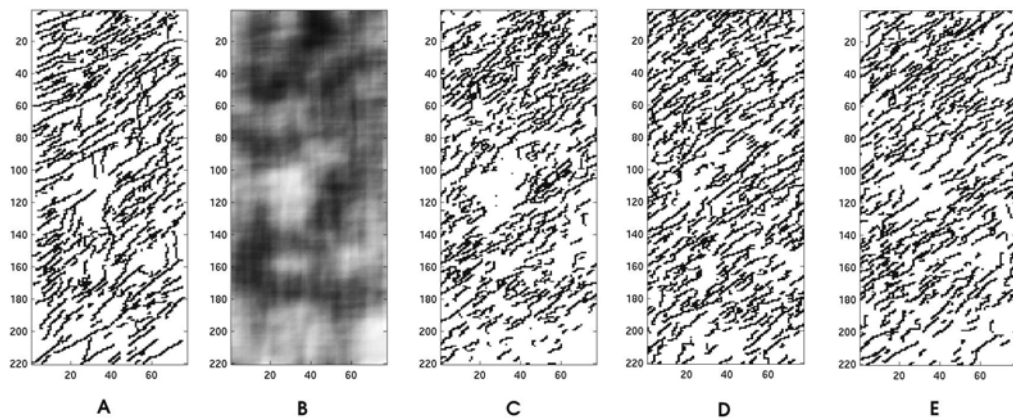
**Fig. 67** (A) Contrainte spatiale ; moyenne mobile de : (B) la réalisation du modèle Tau ; (C) la réalisation de la méthode de classement ; (D) la réalisation de l'approche MPNST.

### 4.3.2 Réseau de fractures

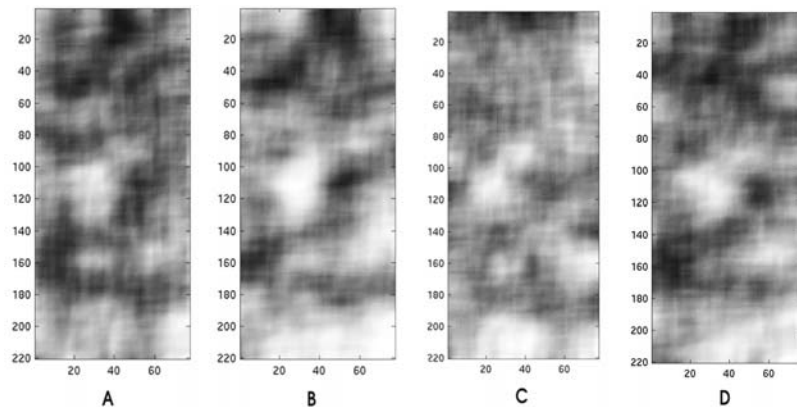
Pour cet exemple, nous avons repris le réseau de fractures utilisé en § 2.3.5 (simulation stationnaire) et en § 3.3.3 (simulation non stationnaire). L'image d'apprentissage principale représente les fractures à 2 dimensions à grande échelle en (100x200 points). La contrainte spatiale fournit l'image d'apprentissage auxiliaire qui représente la densité de fracturation.

La comparaison des trois approches est illustrée (Fig. 68). Retenons ici deux aspects : la reproduction correcte de la géométrie et la restitution de la contrainte spatiale.

Globalement, les trois modèles reproduisent la géométrie de manière satisfaisante. Cependant, sur la réalisation du modèle Tau, il est possible de rencontrer des fractures potentielles tronquées (Fig. 68C). De son côté, la contrainte spatiale (densité de fracturation) n'étant pas très contrastée, il est difficile de juger quelle réalisation la restitue le mieux. Une moyenne locale a été calculée et comparée à la contrainte spatiale (Fig. 69). La réalisation du modèle Tau a visiblement renforcé son impact (Fig. 69C), la valeur adéquate de  $\tau$  reste donc à chercher. La réalisation de la méthode de classement fournit une image plus homogène que la contrainte (Fig. 68D, 69D). L'approche MPNST restitue bien la densité de fracturation sans altérer la structure géométrique (Fig. 68E, 69E).



**Fig. 68 (A) Image d'apprentissage principale ; (B) image d'apprentissage auxiliaire et contrainte spatiale; (C) réalisation du modèle Tau ; (D) réalisation de la méthode de classement ; (E) réalisation de l'approche MPNST.**



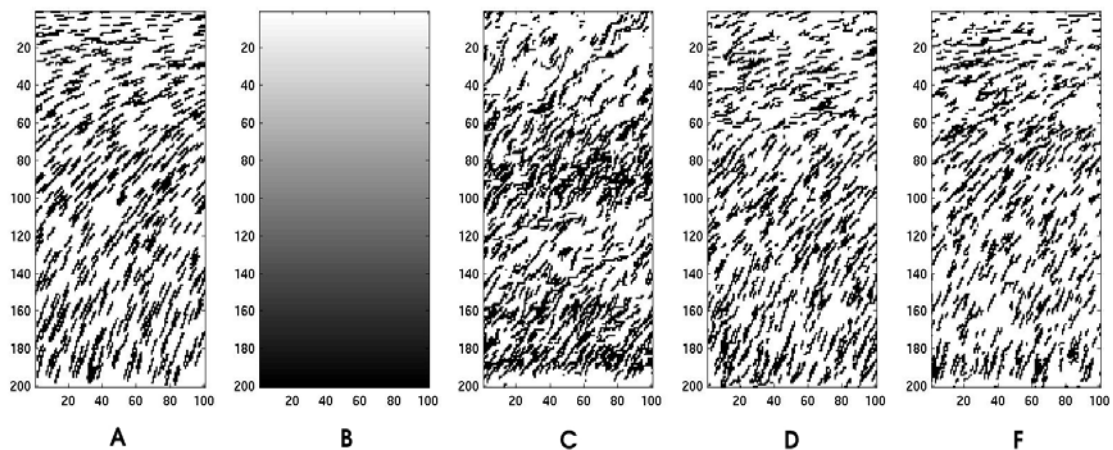
**Fig. 69 Moyenne mobile de : (A) l'image d'apprentissage principale; (B) la réalisation du modèle Tau ; (C) la réalisation de la méthode de classement ; (D) la réalisation de l'approche MPNST.**



### 4.3.3 Segments orientés

Dans les deux exemples précédents, la fonction auxiliaire caractérisait la proportion locale de faciès ; l'exemple de ce paragraphe montre la capacité de l'approche MPNST à prendre en compte des caractéristiques géométriques de nature différente.

Nous avons repris l'exemple utilisé en § 3.3.3. Nous rappelons, que l'image d'apprentissage principale est une réalisation d'un schéma booléen de segments. L'orientation des derniers évolue graduellement de la verticale, en bas de l'image, à l'horizontale, en haut de l'image (Fig. 70A). Cette orientation de segment est utilisée pour créer une image d'apprentissage auxiliaire tout comme un jeu de contraintes (Fig. 70B). Comme notre approche ne dépend pas de la nature de la contrainte spatiale, l'intervalle sur laquelle la fonction auxiliaire est définie n'est pas important. Il faut seulement que la variation de la fonction auxiliaire corresponde à celle de la caractéristique géométrique qui lui est associée. Dans notre exemple, la variation d'orientation des segments du  $0^\circ$  au  $90^\circ$  est représentée par la fonction auxiliaire qui varie entre 0 et 1.



**Fig. 70 (A) Image d'apprentissage principale ; (B) image d'apprentissage auxiliaire et la contrainte spatiale; (C) réalisation du modèle Tau ; (D) réalisation de la méthode de classement ; (E) réalisation de l'approche MPNST.**

Cet exemple illustre bien la capacité de l'approche MPNST à intégrer toute contrainte spatiale, quelque soit sa morphologie.

Avec l'approche MPNST (Fig. 70F), nous pouvons remarquer une raréfaction des segments en haut à droite de la simulation. Ceci est tout à fait justifié compte tenu des configurations similaires observées dans l'image d'apprentissage (Fig. 70A). Voici toutefois quelques simulations fournies par notre approche (Fig. 71), ce afin de dissiper tout doute sur l'éventualité d'un effet de bord.

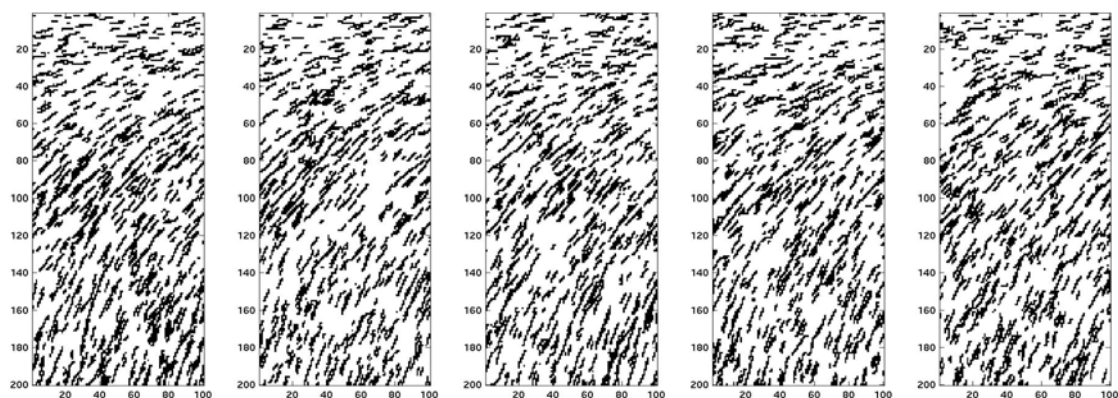


Fig. 71. 5 réalisations de l'approche MPNST.

## 4.4 Résumé

Dans ce chapitre, nous avons proposé un autre algorithme pour intégrer des contraintes non stationnaires dans des simulations multipoints. Par rapport à ce qui se fait traditionnellement, notre algorithme ne repose ni sur un modèle probabiliste particulier ni sur une réduction de variabilité des données auxiliaires. Il consiste à inférer les caractéristiques non stationnaires directement à partir de l'image d'apprentissage tout en gardant la continuité de ces caractéristiques.

Une étude de validation de cette approche a été réalisée sur trois cas : une fonction gaussienne seuillée, un réseau de fractures et un schéma booléen de segments orientés. Les résultats expérimentaux obtenus sont très encourageants. Les exemples simulés par l'approche MPNST reproduisent bien la géométrie tout en respectant, aussi bien globalement que localement, les contraintes. En particulier, l'approche MPNST s'applique à différents types de non stationnarité (la proportion locale des faciès, l'orientation) quelle que soit la nature des données auxiliaires, ce qui n'est pas le cas du modèle Tau.

Pour une utilisation correcte de l'approche MPNST, il suffit d'accorder la variation de la fonction auxiliaire avec celle de la caractéristique géométrique associée. Il convient d'indiquer qu'en utilisant cette approche, seules les structures géométriques qui figurent dans l'image d'apprentissage peuvent être reproduites dans les simulations. Cela montre l'importance du choix de l'image d'apprentissage.

Pour conclure, il faut remarquer que l'approche MPNST est basée sur la simulation MP, ce qui permet de reproduire toute structure géométrique. L'approche MPNST a donc en quelque sorte hérité des avantages ainsi que des inconvénients de l'approche MP. La condition de bonne reproduction des structures géométriques consiste à respecter les limites d'application de l'approche MP (2.3). La condition de reproduction de la contrainte non stationnaire (apportée par l'approche MPNST) consiste à présenter cette non stationnarité sur deux images d'apprentissages. Nous avons pu observer son efficacité dans les exemples traités ci-dessus, car tous ces exemples possèdent les prérequis minima pour la simulation non stationnaire MP :

- l'image d'apprentissage est représentative de la structure géométrique de basse entropie ;
- la taille de l'image d'apprentissage est supérieure à celle de la fenêtre d'exploration qui est à son tour est supérieure à la taille des patterns géométriques répétitifs ;
- la non stationnarité des images d'apprentissage est suffisante pour reproduire la variabilité géométrique imposée par la contrainte spatiale.

## 5 Application au modèle génétique

Dans ce chapitre, nous présentons une application de l'approche MP à un modèle génétique. Ce modèle simule le développement de chenaux au cours du temps, avec l'évolution de méandres, des ruptures de berges, des abandons de chenaux, des formations de dépôts de sable ou d'argile, le tout étant tributaire de l'érodabilité des terrains [Lopez (2003)]. Après avoir décrit le modèle et ses caractéristiques géologiques, nous présentons des résultats de simulation en fonction du degré de raffinement de notre approche.

Tout d'abord, nous testons la version initiale de l'approche MP – code SNESIM [Remy (2003)] – sans intégration de contrainte spatiale. Les résultats obtenus (en 2 et en 3 dimensions) montrent que la structure géométrique complexe en plusieurs faciès de forte hétérogénéité spatiale ne peut pas être reproduite par l'approche MP. Même en simplifiant le modèle par réduction du nombre de faciès, la géométrie continue méandriforme reste difficilement reproductible par l'approche MP. Nous appliquons ensuite l'approche MPNST en commençant par le cas de deux faciès en 2 dimensions. Une fois ce test réussi, nous restons à deux faciès, mais passons en 3D. Finalement, nous abordons la tâche initiale – simulation de 6 faciès en 2 et en 3 dimensions.

### 5.1 Modèle génétique

Le modèle génétique que nous allons introduire ici correspond à la géométrie d'un environnement sédimentaire fluviatile. Ce modèle est décrit par 6 faciès : 1) dépôt résiduel de base de chenal (channel lag) ; 2) bouchon sableux (sand plug) ; 3) barre d'accrétion latérale (point bar) ; 4) levées (levees) ; 5) dépôt de débordement (overbank) ; 6) bouchon argileux (mud plug). La réalisation de ce modèle est effectuée en 3D sur une grille cartésienne de taille  $100 \times 100 \times 1200 = 12\,000\,000$  blocs (Fig. 72). Les dimensions physiques d'un bloc en X, Y et Z sont respectivement de 12,5m, 12,5m et 0,1m.

La réalisation obtenue joue le rôle d'image d'apprentissage pour la simulation MP. Selon les conditions d'application de l'approche MP, nous pouvons décrire cette image de manière suivante :

- 1) La structure est prononcée mais pas répétitive. (Il existe des combinaisons de faciès qui ne se trouvent qu'une seule fois dans l'image d'apprentissage) ;

- 2) L'étendue des objets géologiques est supérieure à la taille de l'image d'apprentissage ;
- 3) Par conséquent, la fenêtre d'exploration ne couvre pas intégralement un objet et ses abords ;
- 4) La variabilité spatiale ne se stabilise pas au-delà d'une distance inférieure à la taille de l'image d'apprentissage. Le trend (changement des proportions ainsi que la géométrie des faciès) le long de la verticale est important (Fig. 73).

Cette image d'apprentissage ne rentre donc pas dans les limites d'application de l'approche MP. Cependant, dans les paragraphes suivants nous allons essayer de la reproduire.

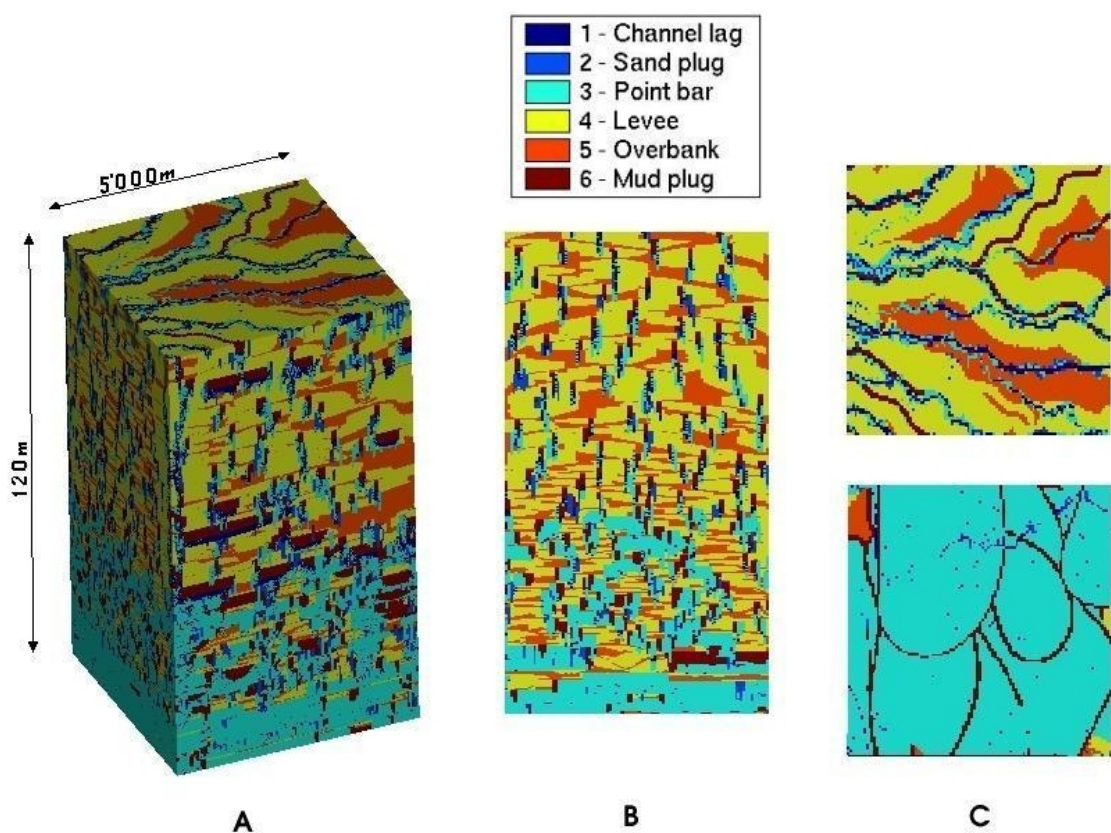
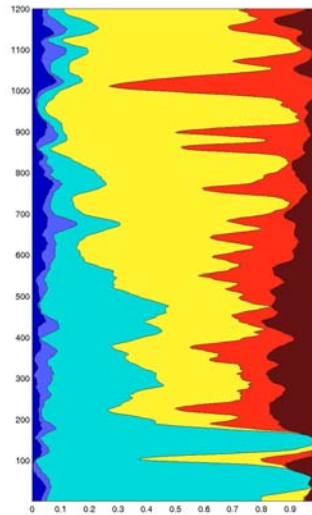


Fig. 72 (A) Réalisation du modèle génétique en 3D, (B) une coupe vertical du modèle; (C) deux couches horizontales qui représentent le toit et la base du modèle.



**Fig. 73** Les courbes de proportions verticales calculées sur l'ensemble du modèle génétique : de gauche à droite les faciès de 1 à 6.

## 5.2 Simulation par l'approche MP

Nous appliquons l'approche MP à la réalisation du modèle génétique. La fenêtre d'exploration utilisée pour la simulation en 3D est présentée sur les Fig. 74-75. Pour donner à l'algorithme la possibilité de mieux reproduire la continuité des chenaux, nous avons choisi un chemin régulier de simulation qui consiste en un balayage de chaque plan horizontal selon leur direction principale. Les résultats des premiers tests pour 6 faciès sont décourageants, comme nous l'avions prévu (Fig. 76).

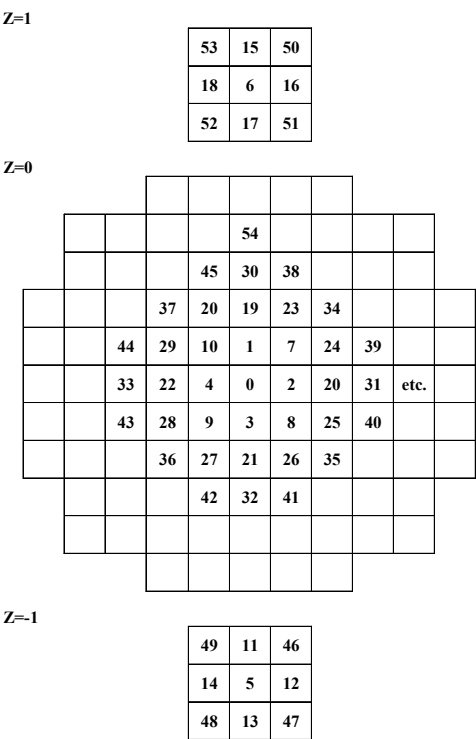


Fig. 74 Fenêtre d'exploration pour la simulation en 3D dans les axes XY avec  $Z = 1$ ,  $Z = 0$  et  $Z = -1$ .

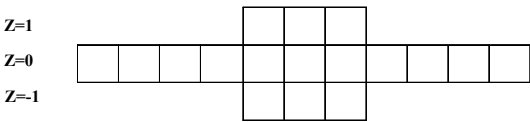
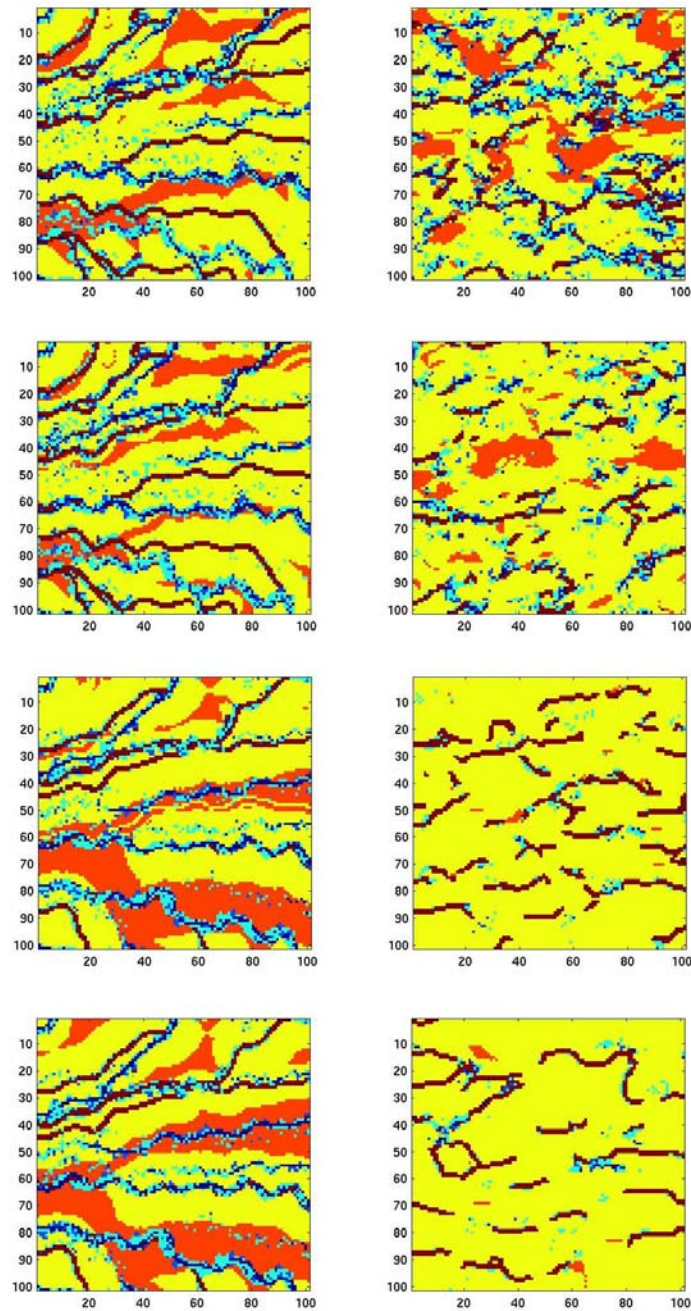


Fig. 75 Fenêtre d'exploration pour la simulation en 3D dans les axes XZ.

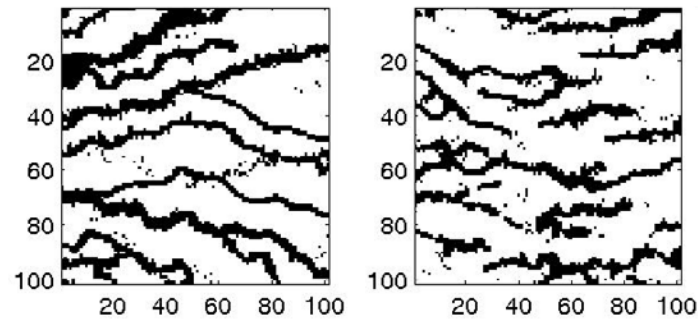


**Fig. 76 Simulation du modèle géologique des 6 faciès en 3D. Quatre couches horizontales de l'image d'apprentissage (à gauche) et les réalisations MP qui leur correspondent (à droite).**

Pour réduire la complexité de la tâche et, plus particulièrement, pour augmenter la répétitivité des motifs géométriques, nous allons tester le cas de la simulation d'une couche horizontale (2D) de géométrie méandrique (deux faciès). Pour obtenir l'image binaire, les faciès 1, 2, 3 et 6 ont été regroupés en faciès noir, et les faciès 4 et 5 — en faciès blanc. Pour améliorer la continuité des chenaux, la simulation est effectuée selon un cheminement régulier qui prend en compte leur orientation. Le résultat de la simulation montre les



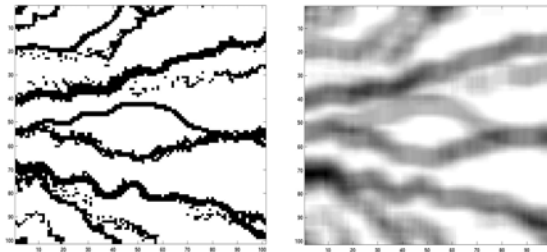
limitations de capacité de l'approche MP à reproduire des objets continus (Fig. 77). Le recours à une post simulation ou bien à une simulation multigrille permettra probablement d'améliorer ces résultats, mais ce n'est pas notre objectif ici.



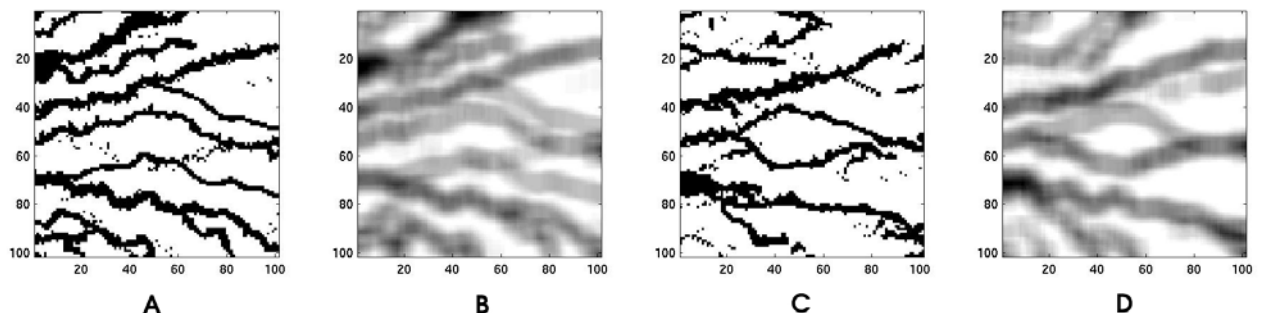
**Fig. 77 Simulation d'une couche horizontale du modèle géologique à 2 faciès : image d'apprentissage à gauche; réalisation MP à droite.**

## 5.3 Simulation par l'approche MPNST

Nous appliquons maintenant l'approche MPNST au modèle géologique en commençant par deux faciès à deux dimensions. Nous allons exploiter la non stationnarité de l'image d'apprentissage pour indiquer les zones de passage les plus vraisemblables des chenaux. Une façon de procéder est de définir l'image des contraintes spatiales comme une moyenne locale d'une image d'apprentissage selon un certain support. Dans ce test, nous avons introduit des contraintes spatiales non pas à partir de l'image d'apprentissage, mais à partir d'une autre image, dite « de référence » (Fig. 78). La simulation est alors effectuée selon un cheminement régulier qui respect le sens principal des chenaux. Le résultat obtenu est maintenant cohérent avec l'image d'apprentissage, sa géométrie continue et la contrainte (Fig. 79).



**Fig. 78** Image de référence à gauche, contrainte spatiale qui en est déduite à droite.



**Fig. 79** Simulation d'une couche horizontale du modèle géologique à 2 faciès : (A) image d'apprentissage principale ; (B) image d'apprentissage auxiliaire ; (C) réalisation MPNST ; (D) contrainte spatiale déduite de l'image de référence.

Nous passons ensuite au cas de deux faciès d'un milieu fluvial à trois dimensions en utilisant la même fenêtre d'exploration qu'en §5.2. Pour pouvoir mieux évaluer les résultats obtenus, la contrainte spatiale a été prise identique à l'image d'apprentissage auxiliaire. Bien que le cas en 3D soit tout à fait comparable au 2D, les résultats sont assez différents (Fig. 81). Nous nous trouvons confronté au problème de l'ordre de conditionnement aux points voisins. L'ordre de voisinage dans la fenêtre d'exploration en

3D définit la priorité selon laquelle le conditionnement s'effectue : priorité à la verticale ou bien à l'horizontale (Fig. 80). En donnant la priorité à la verticale (Fig. 80A), on perd la continuité horizontale et inversement (Fig. 80B). Ce problème nous amène à un compromis dans l'ordre de voisinage, consistant à faire alterner le conditionnement horizontal et vertical comme le montre la Fig. 74.

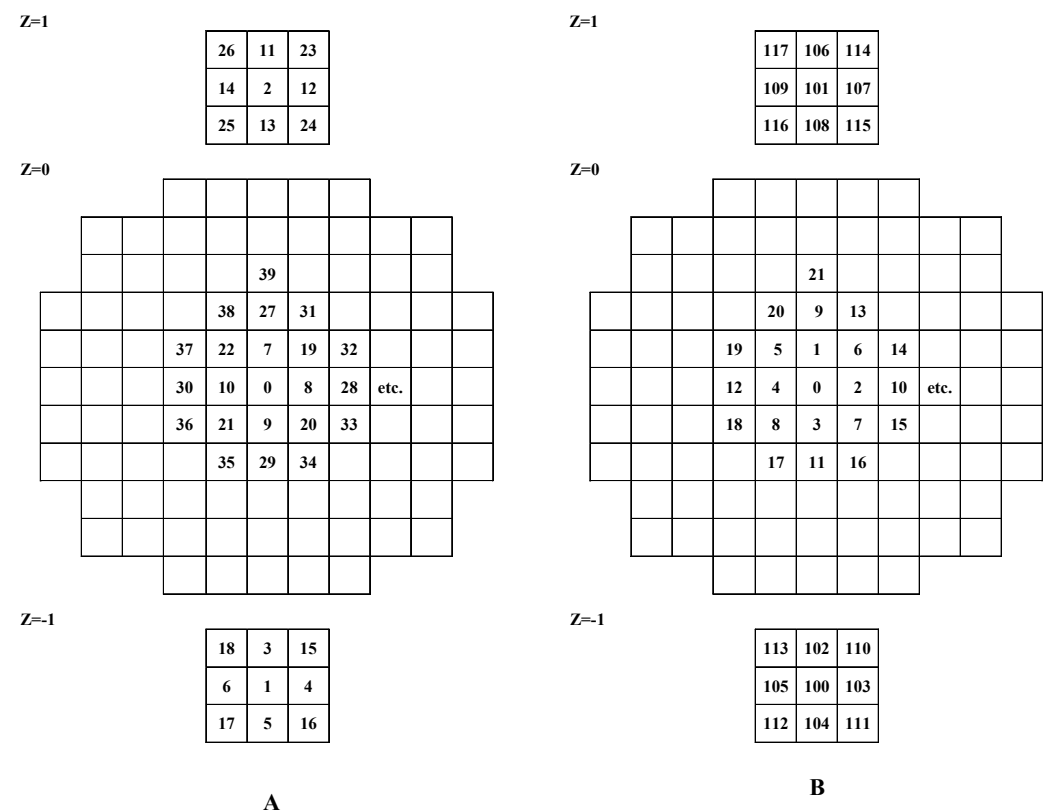
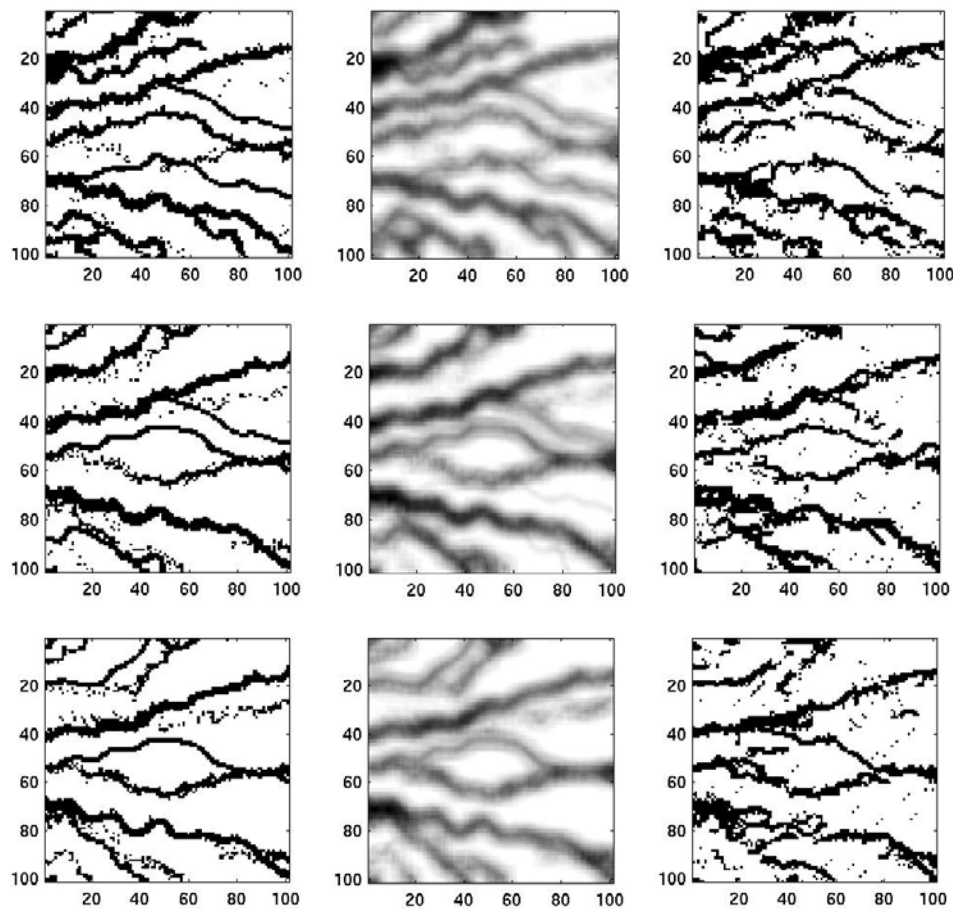


Fig. 80 Fenêtres d'exploration pour la simulation en 3D dans les axes de XY avec Z =1, Z=0 et Z=-1. L'ordre de conditionnement s'effectue : (A) d'abord par des points voisins verticaux ; (B) d'abord par des points voisins horizontaux.

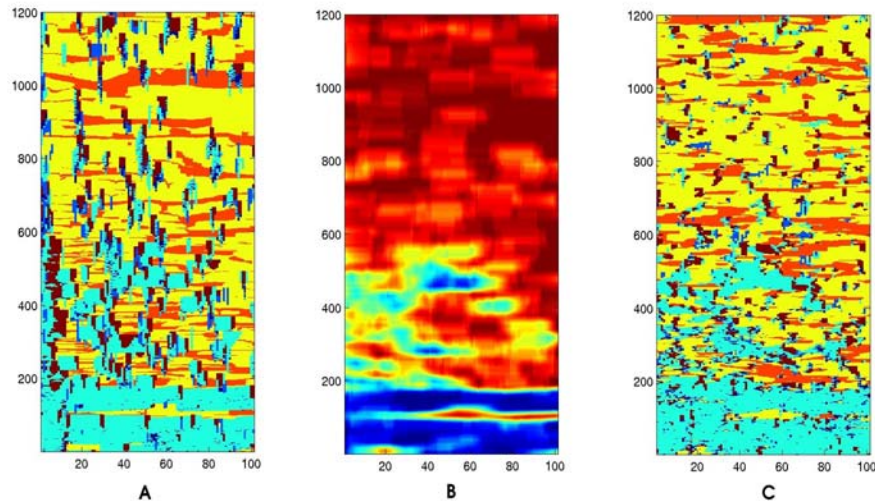


**Fig. 81 Simulation du modèle géologique à 2 faciès en 3D. Les trois couches horizontales de : l'image d'apprentissage principale (à gauche) ; l'image d'apprentissage auxiliaire et la contrainte spatiale (au centre); la réalisation MPNST (à droite).**

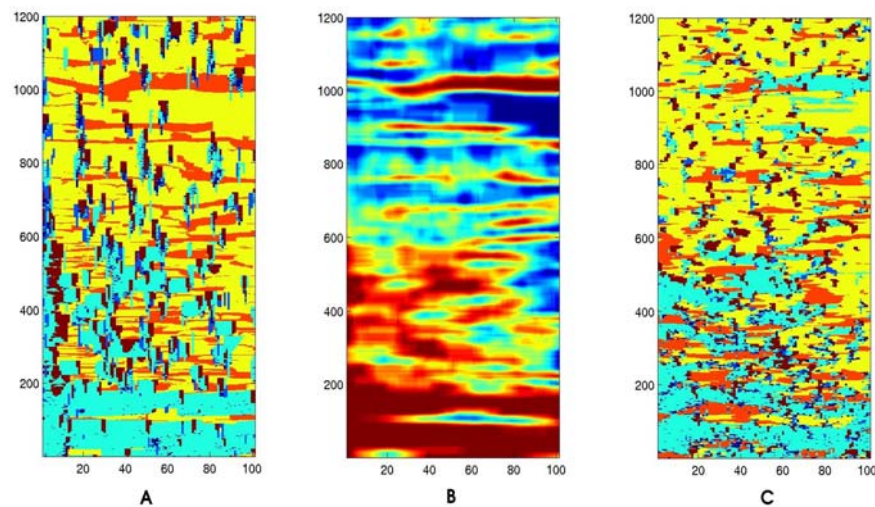
Nous abordons finalement la reproduction du modèle géologique à 6 faciès. Nous commençons par le cas en 2D. L'image d'apprentissage représente une coupe verticale du modèle (Fig. 82A). Pour effectuer une simulation à plus de 2 faciès nous proposons deux approches : la première repose sur la simulation simultanée des 6 faciès en les conditionnant à la contrainte de l'un des faciès ; la deuxième consiste à simuler indépendamment chaque faciès avec sa propre contrainte puis à mettre en commun les résultats obtenus.

Nous testons tout d'abord la première approche. Les proportions du faciès 3 (barre d'accrétion latérale) et du faciès 4 (levées) sont dominantes dans le modèle géologique. Elles définissent l'hétérogénéité verticale du modèle de manière complémentaire. Aussi avons nous choisi de conditionner la simulation par l'une de ces proportions. La simulation simultanée des 6 faciès donne des résultats médiocres (Fig. 82-83). La structure géométrique et les contacts entre faciès ressemblent à ceux de l'image d'apprentissage. Les proportions locales du faciès contraint sont bien restituées. Par contre, les proportions

locales des autres faciès sont libres de toute contrainte. Les faciès sont donc placés au hasard. Une telle approche ne présente d'intérêt que dans les cas où nous cherchons à représenter la variabilité maximale du modèle (e.g. estimation d'incertitude) et dans les cas où les proportions des tous les faciès ne sont pas connues.



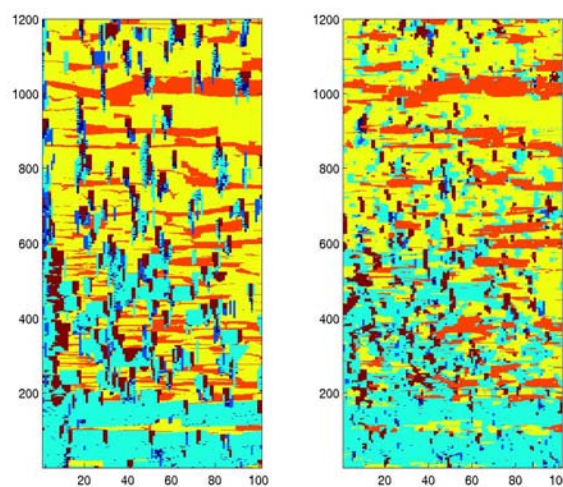
**Fig. 82 Simulation d'une coupe verticale de 6 faciès en 2D : (A) image d'apprentissage principale; (B) image d'apprentissage auxiliaire et contrainte (la proportion locale du faciès 3) ; (C) réalisation MPNST.**



**Fig. 83 Simulation d'une coupe verticale de 6 faciès en 2D : (A) image d'apprentissage principale; (B) image d'apprentissage auxiliaire et contrainte (la proportion locale du faciès 4) ; (C) réalisation MPNST.**

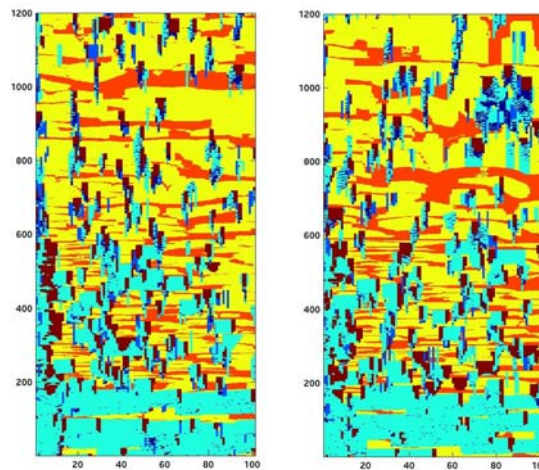
Nous en venons maintenant au test de deuxième approche qui consiste à travailler faciès par faciès. Pour pouvoir les recomposer, nous utilisons des contraintes spatiales (proportions locales de chaque faciès) déduites de l'image d'apprentissage. Même après avoir contraint chaque faciès à sa proportion locale, il demeure des fluctuations statistiques qui font que différents faciès peuvent se chevaucher. Le résultat final fourni par la

recomposition des 6 faciès simulés dépend donc de l'ordre dans lequel ils sont pris en compte. Le choix de cet ordre doit être en accord avec la géologie et la géométrie du modèle à reproduire. Cela veut dire, que les faciès de granulométrie fine et très fine (déposés dans des environnements de basse énergie) peuvent être érodés par les faciès de granulométrie plus grossière (déposés dans des environnements de haute énergie). Ainsi, le faciès formé dans des environnements de plus haute énergie a la priorité par rapport aux autres. Dans notre cas, le modèle exige l'ordre suivant de faciès : 1, 2, 6, 5, 4, 3 (par exemple, si les faciès 6 et 5 sont présents au même point, on garde le faciès 6). Le résultat obtenu par cette approche ressemble à la structure géométrique de l'image d'apprentissage et restitue les proportions locales des 6 faciès (Fig. 84).

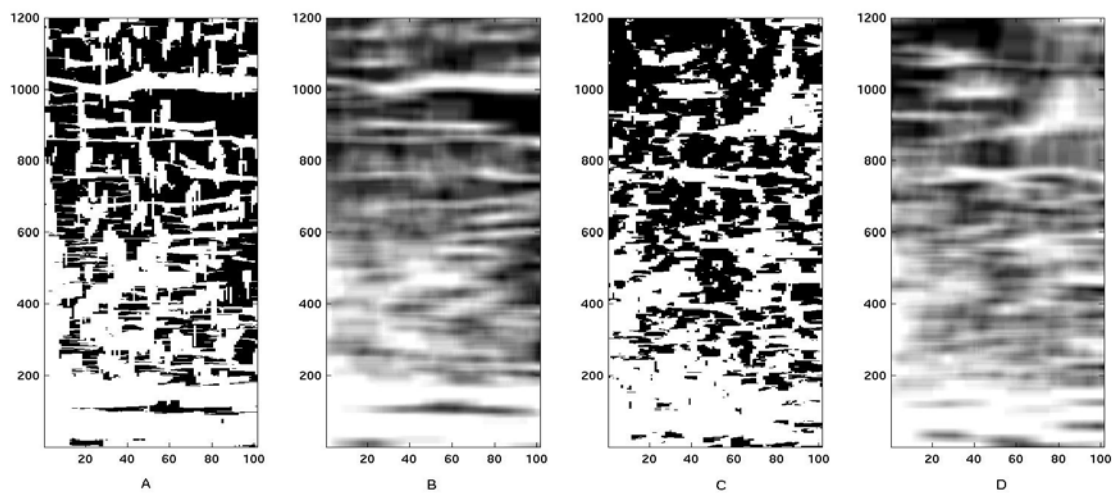


**Fig. 84 Simulation d'une coupe verticale du modèle géologique à 6 faciès en 2D : image d'apprentissage (à gauche) ; reconstitution des 6 faciès simulés par MPNST (à droite).**

Nous envisageons ici le cas plus complexe où les contraintes spatiales sont différentes des images d'apprentissage auxiliaires. Lorsque les dernières sont déduites (tout comme dans le test précédant) de l'image d'apprentissage (Fig. 85A), les premières sont déduites d'une autre image, dite « de référence » (Fig. 85B). Les Fig. 86-87 montrent quelques résultats intermédiaires de simulation. Les 6 faciès simulés sont mis en commun selon le même ordonnancement. Le résultat final de ce test montre que la structure géométrique obtenue ressemble à celle de l'image d'apprentissage et restitue bien les proportions locales de faciès de l'image de référence (Fig. 84).

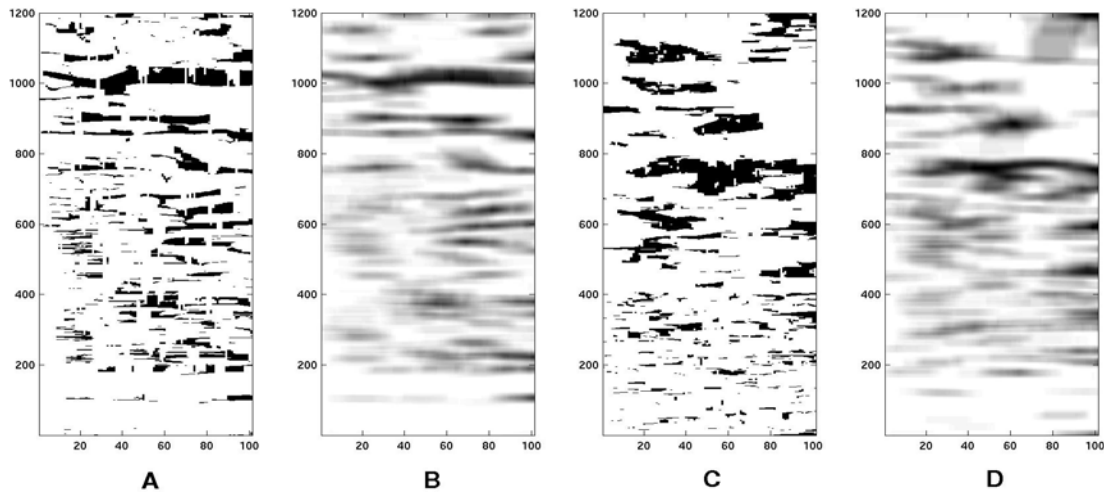


**Fig. 85** Image d'apprentissage (à gauche) est une coupe verticale de la réalisation du modèle géologique. Image de référence (à droite) est une coupe verticale d'une autre réalisation du modèle géologique.

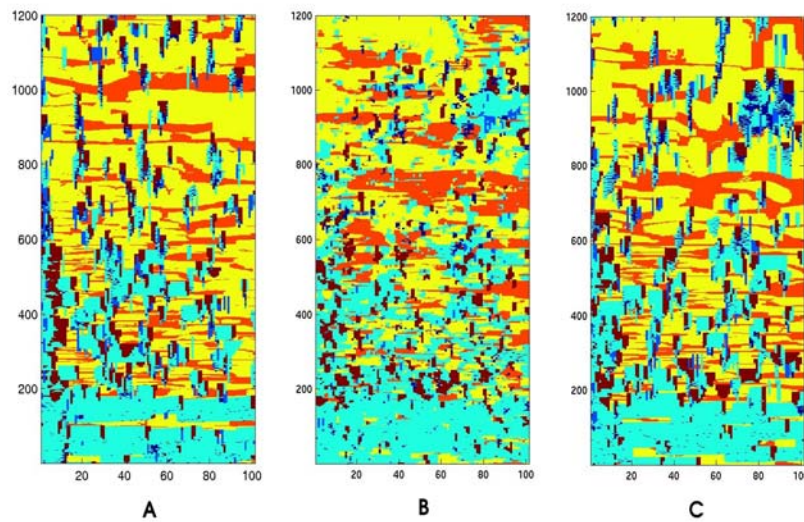


**Fig. 86** Simulation d'une coupe verticale du faciès 4 en 2D : (A) image d'apprentissage principale; (B) image d'apprentissage auxiliaire ; (C) réalisation MPNST ; (D) contrainte déduite de l'image de référence.





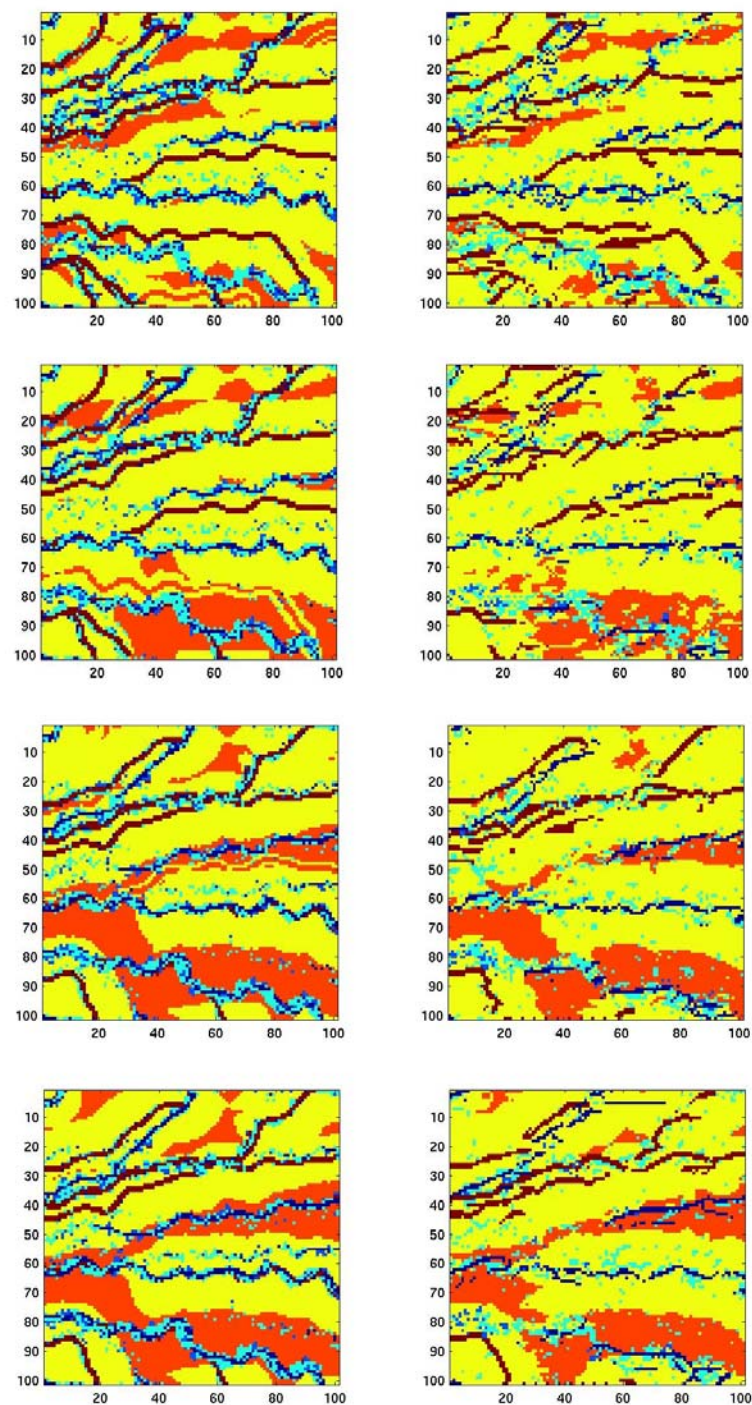
**Fig. 87** Simulation d'une coupe verticale du faciès 5 en 2D : (A) image d'apprentissage principale; (B) image d'apprentissage auxiliaire ; (C) réalisation MPNST ; (D) contrainte spatiale déduite de l'image de référence.



**Fig. 88** Simulation d'une coupe verticale du modèle géologique à 6 faciès en 2D : (A) image d'apprentissage ; (B) recombinaison des 6 faciès simulés par MPNST; (C) image de référence.

Maintenant convaincus par la capacité de notre approche à reproduire la non stationnarité spatiale tout en préservant la structure géométrique, nous abordons le cas à 6 faciès à 3D. En répétant toute le travail faciès par faciès, leur simulation sous contrainte de la proportion locale et leur assemblage, nous obtenons la simulation par MPNST (Fig. 89).





**Fig. 89** Simulation du modèle géologique à 6 faciès à 3D. Quatre couches horizontales de l'image d'apprentissage (à gauche) et les réalisations MPNST qui leur correspondent (à droite).

## 5.4 Résumé

Nous venons de voir les résultats de simulation d'un modèle génétique par l'approche MP. Parmi eux, certains sont loin d'être satisfaisants et montrent que l'approche MP est loin d'être une panacée.

Commençons par les simulations incorrectes. Ce sont les simulations stationnaires d'une géométrie continue (fluviale), multifaciès à 3D. Les raisons de cet échec sont les suivantes :

- 1) la longueur des chenaux est supérieure à la taille de l'image d'apprentissage, ainsi qu'à la taille de fenêtre d'exploration ;
- 2) le manque de répétitivité de patterns réduit le nombre de points voisins effectivement utilisés pendant la simulation. L'augmentation de la fenêtre d'exploration ne résoudra pas le problème (§ 2.3.2);

La réduction du nombre de faciès a amélioré la reproduction de la géométrie des chenaux. Avec deux faciès, la répétitivité de patterns est plus forte.

La simulation :

- 1) de 2 faciès à 2D,
- 2) selon l'approche MPNST (en intégrant la contrainte spatiale),
- 3) selon un cheminement régulier,

permet de reproduire la continuité de la structure géométrique. Le même exemple à 3D amène des difficultés liées au choix de l'ordre de voisinage dans la fenêtre d'exploration.

La simulation de 6 faciès à 2D (à condition de respecter les contacts entre faciès) est possible selon deux approches. La première consiste à simuler les 6 faciès simultanément sous la contrainte du faciès dominant. Cette approche est intéressante dans la mesure où elle fournit des réalisations variées où les faciès sont distribués de façon aléatoire. La seconde approche consiste à simuler chaque faciès indépendamment. La mise en commun des 6 faciès simulés est possible en les contraignant selon leurs proportions locales par l'approche MPNST. L'ordre de la recombinaison des faciès lors de la mise en commun doit être spécifié en accord avec le modèle géologique.

En résumé, l'approche MP n'est pas convaincante lorsqu'elle est utilisée pour reproduire des structures géométriques continues. Elle n'est applicable que sous certaines conditions ; et dès que l'une d'entre elles n'est pas satisfaite, les résultats obtenus peuvent être de qualité douteuse. Dans une telle situation, l'ajustement des paramètres d'entrée (la configuration de la fenêtre d'exploration, l'ordre de voisinage dans la fenêtre

d'exploration, le choix de cheminement etc.) prend un rôle décisif pour l'amélioration du résultat [Maharaja (2004b)].

Appliquée à un modèle génétique, l'approche MPNST permet de reproduire correctement la structure géométrique tout en respectant les contraintes spatiales. Plus la structure (multiples faciès, objets continus, 3D) est complexe, plus le choix des paramètres de simulation (l'ordre de voisinage dans la fenêtre d'exploration, chemin prédéfinie ...) est délicat. Notons toutefois que le choix des paramètres ne fait pas tout. Utilisés conjointement dans les approches MP et MPNST, pour reproduire des géométries complexes de forte hétérogénéité, seule la deuxième approche fournira des résultats corrects. C'est parce qu'elle seule est capable de prendre en compte des contraintes auxiliaires.

## 6 Conclusions et perspectives

Cette thèse est consacrée à l'approche multipoints et son application à la simulation géostatistique. En raison de son utilisation intensive, il nous a semblé utile d'effectuer une analyse approfondie de cette approche. Cela permet de mieux cerner ses avantages et ses inconvénients, ainsi que ses limites d'application.

Initialement développée sous hypothèse de stationnarité, l'approche MP ne permettait pas au départ de simuler de structure géométrique de fort trend spatial. D'autres approches sont ensuite apparues pour pallier cette absence de stationnarité via l'intégration d'une information auxiliaire. Deux d'entre elles ont été examinées, à savoir le modèle Tau et la méthode de classement. Leurs inconvénients nous ont amené à envisager une autre approche.

### 6.1 Approches MP

L'approche MP se montre très efficace pour simuler des géométries complexes si les conditions suivantes d'application sont respectées :

- 1) L'image d'apprentissage contient des patterns répétitifs (de faible entropie) ;
- 2) La taille de cette image est suffisamment grande par rapport à celle de la fenêtre d'exploration ;
- 3) La taille de la fenêtre d'exploration est suffisamment grande par rapport à celle des patterns répétitifs de l'image d'apprentissage ;
- 4) La fonction aléatoire  $Z(\cdot)$  est supposée stationnaire de portée limitée, i.e. la variabilité spatiale de l'hétérogénéité présenté dans l'image d'apprentissage se stabilise au-delà d'une certaine distance bien inférieure à la taille de l'image.

Hormis ces conditions, les paramètres d'entrée jouent un rôle déterminant sur le résultat de simulation. Pour améliorer la reproduction d'une structure géométrique, nous conseillons :

- 1) d'augmenter la taille de l'image d'apprentissage ou bien d'utiliser de

multiples images d'apprentissage;

- 2) d'augmenter la taille de la fenêtre d'exploration (tout en étant conscient que si la structure géométrique n'est pas répétitive, le nombre de points voisins effectivement utilisés reste assez limité);
- 3) d'adapter la configuration et l'ordre de voisinage de fenêtre d'exploration à la géométrie des patterns répétitifs;
- 4) d'utiliser le chemin de simulation orienté dans le sens principal de la continuité des objets pour pouvoir mieux les reproduire ;
- 5) d'utiliser la technique multigrille pour mieux reproduire des objets de taille supérieure à celle de la fenêtre d'exploration ;
- 6) d'utiliser la technique de « post-processus » pour affiner la reproduction de géométrie complexe.

Même en suivant ces conseils d'ajustement des paramètres d'entrée, l'utilisation actuelle de l'approche MP n'est pas convaincante pour reproduire la complexité des structures multifaciées que l'on rencontre en pratique.

## 6.2 Approches existantes

Les deux approches existantes de simulation non stationnaire sont de nature différente. Nous les résumons, donc, séparément.

L'analyse théorique du modèle Tau et sa validation sur trois modèles de probabilités soulignent, que l'approximation de la probabilité biconditionnelle par une fonction de probabilité monoconditionnelle n'est correcte que pour quelques cas particuliers.

Au vu des simulations obtenues à l'aide du modèle Tau, on peut conclure que ce modèle n'est vraiment utilisable que lorsque la valeur du paramètre  $\tau$  est explicitement connue. Cela n'est jamais le cas en pratique. De plus, la non stationnarité reproduite à l'aide du modèle Tau ne porte que sur la proportion locale de faciès. La non stationnarité de caractéristiques géométriques, telles que l'orientation d'objets, la taille, la forme etc., ne peut pas être reproduite par ce modèle. Il faut également prendre en compte que la simulation par le modèle Tau ne fournit qu'un compromis entre une bonne reproduction de la géométrie et un bon respect des contraintes.

La méthode de classement semble plus adaptée à la simulation non stationnaire. Elle repose sur l'extraction de probabilités empiriques biconditionnelles à partir de l'image

d'apprentissage, mais elle réduit la variété des données auxiliaires (souvent continues) à seulement quelques classes, ce qui peut engendrer des artefacts dans les simulations.

## 6.3 Approche MPNST

L'approche MPNST a été élaborée pour éviter les inconvénients des deux approches précédentes. Elle possède les caractéristiques suivantes :

- 1) indépendance de modèles probabilistes analytiques ;
- 2) conditionnement par des données auxiliaires représentant la non stationnarité ;
- 3) pas de simplification (classement) des données auxiliaires ;
- 4) indépendance de la nature des données auxiliaires ;
- 5) bonne reproduction de la géométrie conditionnée par la contrainte spatiale.

L'approche MPNST consiste à réaliser deux tâches : reproduire la géométrie et la conditionner selon la contrainte spatiale [Chugunova et Hu (2006), Chugunova et Hu (2008)]. Le respect de la contrainte spatiale dépend de la représentativité des images d'apprentissage principale et auxiliaire. La reproduction de la géométrie souffre des mêmes conditions limites que l'approche MP. En pratique toutefois, l'approche MPNST surpasse les limites de fonctionnement de MP et permet de reproduire effectivement des géométries complexes multifaciées non stationnaires [Chugunova et al. (2007)].

Un logiciel prototype (non commercial) a été développé au sein de l'Institut Français du Pétrole. Ce logiciel effectue la simulation MPNST en 2D et 3D pour plusieurs faciès en intégrant des contraintes auxiliaires continues. L'algorithme MPNST permet en effet d'élaborer un logiciel qui intégrera plus d'un jeu de contraintes auxiliaires continues. Cependant, plus les contraintes sont nombreuses, plus il est difficile de trouver les répliques qui les respectent. Théoriquement, pour assurer la bonne reproduction d'une structure géométrique par l'approche MPNST multi-contrainte, il suffit que l'ensemble des gammes de valeurs prises par les fonctions auxiliaires soit bien représenté dans l'image d'apprentissage.

## 6.4 Application au modèle génétique

Les approches MP et MPNST ont été appliquées à un modèle génétique. Les résultats de simulation 2D à deux faciès montrent que l'approche MP ne reproduit pas correctement les objets continus tandis que l'approche MPNST, de par son aptitude à intégrer des contraintes auxiliaires, permet cette reproduction. Les résultats fournis par la simulation MP en 3D et sur 6 faciès sont proprement décourageants. Ceux de l'approche MPNST sont nettement plus corrects : la géométrie simulée ressemble à celle de l'image d'apprentissage – en particulier la continuité des chenaux est préservée – tout en respectant les contraintes.

## 6.5 Perspectives

De nos jours, l'approche MP connaît de nombreuses applications [Zhang et al. (2007), Suzuki et al. (2007), Gonzalez et al. (2007)]. Elle évolue dans plusieurs directions, telles que la reproduction des structures géométriques par pattern [Eskandari et Srinivasan (2007)], l'amélioration du code SNESIM [Suzuki et Strebelle (2007), Daly et Knudby (2007)], l'analyse d'incertitude et la valorisation de gisement en appliquant MP aux différents scénarios géologiques [Scheidt et Caers (2007)] etc. Mais, cette évolution est toujours limitée par une hypothèse de stationnarité de l'image d'apprentissage. L'utilisation de l'approche MPNST à la place de MP stationnaire élargit ces directions. Mais d'autres perspectives peuvent être envisagées, telles que :

- 1) l'élaboration d'un logiciel pour effectuer des simulations MPNST multi-contraintes ;
- 2) l'intégration de l'approche MPNST dans un modèleur numérique (par exemple gOcad) pour obtenir de multiples réalisations et pouvoir estimer des incertitudes de production à l'aide d'un modèle plus réaliste que les modèles actuellement en cours (fonctionnelles gaussiennes seuillées ou modèles de type objets) ;
- 3) la mise au point d'un algorithme pour transiter de façon continue entre les réalisations de l'approche MPNST, par exemple par déformation graduelle [Hu (2000)]. Ceci devrait faciliter le calage hydrodynamique des données [Hoffman et Caers (2004), Kashib et Srinivasan (2005), Caers (2006)], indispensable en dernière étape de modélisation numérique d'un réservoir pétrolier ;

- 4) la mise en œuvre opératoire de l'approche MPNST à des cas réels de géométrie complexe et hétérogène. Nous sommes convaincus que l'élaboration de nouvelles approches (méthodes, algorithmes) doit répondre d'abord à des nécessités pragmatiques, puis éventuellement à des ambitions théoriques.



## 7 Bibliographie

Akamine-Ramirez, J. and Caers, J. (2006) Conditioning stochastic realisation to hard data with varying reliability. In *Proc. 11th International Congress, University of Liege, Belgium, International Association for Mathematical Geology*.

Allard, D. (1994) Simulating a geological lithofacies with respect to connectivity information using the truncated gaussian model. *Geostatistical Simulations*, 197-211. Netherlands, Kluwer Academic.

Arpat, B. and Caers, J. (2004a) A multiple-scale, pattern-based approach to sequential simulation. O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 255-264. Netherlands, Springer.

Arpat, B. and Caers, J. (2004b) Reservoir characterization using multiple-scale geological patterns. In *Proc. 9th European Conference on the Mathematics of Oil Recovery*.

Arpat, B., Strebel, S., and Caers, J. (2002) Feature-based geostatistics: an application to a submarine channel reservoir. In *Proc. SPE 77426*. San Antonio, Texas, SPE Annual Technical Conference and Exhibition.

Barrera, A. E., Ni, J., and Srinivasan, S. (2004) Multiple point geostatistics: optimal template selection and implementation in multi-thread computational environments. O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 621-630. Netherlands, Springer.

Baudrier, E. (2005) Comparaison d'images binaires reposant sur une mesure locale des dissimilarités. Application à la classification. PhD thesis, Université de Reims Champagne-Ardenne, UFR des Sciences Exactes et Naturelles.

Beucher, H. (1998) Intégration d'information sismiques pour la simulation de réservoirs. PhD thesis, École Nationale Supérieure des Mines de Paris.

Beucher, H., Benito Garcia-Morales, M., and Geffroy, F. (2004) Fitting the boolean parameters in a non-stationary case. O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 569-578. Netherlands, Springer.

Bour, O., Davy, P., Darcel, C., and Odling, N. (2002) A statistical scaling model for fracture network geometry with validation on a multiscale mapping of a joint network. (Hornelen Basin, Norway) *Journal of Geophysical Research* 107[No 0], 2001-2004. Géoscience Rennes, CNRS UMR 6118.

Bour, O., Davy, P., and de Dreuzay, J.-R. (2001) Hydraulic properties of two-dimensional random fracture networks following a power law fault length distribution.

---

1. Permeability of networks based on lognormal distribution of apertures. 2001WR900010. Water Resources Research 37[No 8], 2079-2095. Géoscience Rennes, UMR 6118, CNRS.

Cacas, M. C., Daniel, J. M., and Letouzey, J. (2001) Nested geological modelling of naturally fractured reservoirs. *Petroleum Geoscience* 7, S43-S52.

Caers, J. (2006) The probability perturbation method: a new look at Bayesian inverse modeling. *Math.Geology* in press.

Caers, J. and Journel, A. G. (1998) Stochastic reservoir simulation using neural networks trained on outcrop data. In *Proc. SPE* 49026. New Orleans, SPE Annual Technical Conference and Exhibition.

Caers, J., Srinivasan, S., and Journel, A. G. (1999) Geostatistical quantification of geological information for a fluvial-type North Sea reservoir. In *Proc. SPE* 56655. Houston, SPE Annual Technical Conference and Exhibition.

Caers, J. and Zhang, T. (2002) Multiple-point geostatistics: a quantitative vehicle for integrating geologic analog into multiple reservoir models. Rapport. Stanford Center for Reservoir Forecasting.

Caumon, G. and Journel, A. G. (2006) Early Uncertainty assessment: Application to a hydrocarbon reservoir appraisal.

Caumon, G., Strebelle, S., Caers, J. K., and Journel, A. G. (2004) Assessment of global uncertainty for early appraisal of hydrocarbon fields. In *Proc. Huston, Texas, SPE Annual Technical Conference and Exhibition*.

Chauvet, P. (1999) Aide-mémoire de géostatistique linéaire. Paris, Grou-Radenez.

Chilès, J. P. and Delfiner, P. (1999) Geostatistics : modeling spatial uncertainty. Wiley, New York.

Chugunova, T. and Hu, L.-Y. (2006) Multiple-point statistical simulations constrained by continuous auxiliary data. In *Proc. 11th International Congress, Université de Liège, Belgium, International Association for Mathematical Geology*.

Chugunova, T. and Hu, L.-Y. (2008) Multiple-point simulations constrained to continuous auxiliary data. *Mathematical Geosciences*, Vol. 40, No.2.

Chugunova, T., Hu, L.-Y., and Lerat, O. (2007) Conditioning a process-based fluvial model using a non-stationary multiple-point statistics approach. In *Proc. Petroleum Geostatistics, Cascais, Portugal, European Association of Geoscientists & Engineers*.

---

Daly, C. (2004) Higher order models using entropy, markov random fields and sequential simulation. O.Leuangthong and C.V.Deutsch(eds.), Geostatistics Banff, 215-224. Netherlands, Springer.

Daly, C. and Knudby, C. (2007) Multipoint statistics in reservoir modelling and in computer vision In *Proc. Petroleum Geostatistics, Cascais, Portugal, European Association of Geoscientists & Engineers*.

De Dreuzy, J.-R., Darcel, C., Davy, P., and Bour, O. (2004) Influence of spatial correlation of fracture centers on the permeability of two-dimensional fracture networks following a power law distribution. *Water Resource Research* 40.

Deutsch, C. and Wang, L. (1996) Hierarchical object-based stochastic modeling of fluvial reservoirs. *Math.Geology* 28[No 7], 857-880.

Dubrule, O. (1998) A review of stochastic models for petroleum reservoirs. In *Proc. BSRG meeting on quantification of sediment body geometries and their internal heterogeneities*. London.

Eskandari, K. and Srinivasan, S. (2007) Growthsim - a multiple point framework for pattern simulation. In *Proc. Petroleum Geostatistics, Cascais, Portugal, European Association of Geoscientists & Engineers*.

Feyen, L. and Caers, J. (2004) Multiple-point geostatistics: a powerful tool to improve groundwater flow and transport prediction in multi-modal formations. In *Proc. Neuchatel, Switzerland, GeoEnv*.

Garcia-Morales, M. B. (2003) Non stationnarité dans les modèles de type booléen: application à la simulation d'unités sédimentaires. PhD thesis, École Nationale Supérieure des Mines de Paris.

Gauthier, B. D. M., Garcia, M., and Daniel, J.-M. (2002) Integrated fractured reservoir characterization: a case study in a North Africa field. In *Proc. SPE 65118. Paris, SPE European Petroleum Conference*.

Gonzalez, E. F., Mavko, G., O'Halloran, A., and Gesbert, S. (2007) Seismic inversion using multi-point geostatistics and rock physics. In *Proc. Petroleum Geostatistics, Cascais, Portugal, European Association of Geoscientists & Engineers*.

Guardiano, F. and Strivastava, R. M. (1993) Multivariate geostatistics: beyond bivariate moments. Soares, A.(eds.), Geostatistics-Troia, [1], 133-144. Dordrecht, Kluwer Academic.

Harding, A., Strebelle, S., Levy, M., Thorne, J., Xie, D., Leigh, S., Preece, R., and Scamman, R. (2004) Reservoir facies modelling: new advances in MPS. O.Leuangthong and C.V.Deutsch(eds.), Geostatistics Banff, 559-568. Netherlands, Springer.

---

Hoffman, B. T. and Caers, J. (2004) History matching with the probability perturbation method: application to a North Sea reservoir. In *Proc. Cannes, France, 9th European Conference on the Mathematics of Oil Recovery*.

Hu, L.-Y. (2000) Gradual deformation and iterative calibration of gaussian-related stochastic models. *Math.Geology* 32[No 1], 87-108.

Hu, L.-Y. (2002) Modélisation stochastique des réservoirs hétérogènes. Mémoire pour obtenir l'habilitation à diriger des recherches. PhD thesis, Université Pierre et Marie Curie (Paris VI).

Hu, L.-Y. (2003) History matching of object-based stochastic reservoir models. *SPE Journal*[81503].

Hu, L.-Y. and Le Ravalec-Dupin, M. (2004) On some controversial issues of geostatistical simulation. O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 175-184. Netherlands, Springer.

Jenni, S. (2005) Calage de la géométrie des réseaux de fractures aux données hydrodynamiques de production d'un champ pétrolier. PhD thesis, Université Pierre et Marie Curie (Paris VI).

Jenni, S., Hu, L.-Y., Basquet, R., and Bourbiaux, B. (2004) History matching of stochastic models of field-scale fractures: Methodology and case study. In *Proc. SPE 90020*. Huston, Texas, SPE Annual Technical Conference and Exhibition.

Journal, A. G. (2002) Combining knowledge from diverse sources: an alternative to traditional data independence hypotheses. *Math.Geology* 34[No 5], 573-596.

Journal, A. G. (2004) Beyond covariance: the advent of multiple-point geostatistics. O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 225-233. Netherlands, Springer.

Journal, A. G. and Alabert, F. G. (1990) New method for reservoir mapping. *Journal of Petroleum Technology*, February, 212-223.

Kashib, T. and Srinivasan, S. (2005) A probabilistic approach to integrating dynamic data in reservoir models. *Petroleum Science and Engineering* 50, 241-257.

Kazemi, H., Atan, S., Al-Matrook, M., Dreier, J., and Ozkan, E. (2005) Multilevel fracture network modelling of naturally fractured reservoirs. *SPE Journal*[93053].

Koltermann, C. and Gorelick, S. (1996) Heterogeneity in sedimentary deposits: a review of structure-imitating, process-imitating, and descriptive approaches. *Water Resource Research* 32, 2617-2658.

---

Krishnan, S. (2004) Experimental study of multiple-support, multiple-point dependence and its modeling. O.Leuangthong and C.V.Deutsch(eds.), Geostatistics Banff, 1109-1115. Netherlands, Springer.

Krishnan, S., Boucher, A., and Journel, A. G. (2004) Evaluating information redundancy through the TAU model. O.Leuangthong and C.V.Deutsch(eds.), Geostatistics Banff, 1037-1046. Netherlands, Springer.

Krishnan, S. and Journel, A. G. (2003) Spatial connectivity: from variograms to multiple-point measures. Math.Geology 35[No 8 Nov].

Lantuejoul, C. (2002) Geostatistical simulation, models and algorithms. Springer-Verlag, Berlin.

Le Ravalec-Dupin, M. and Hu, L.-Y. (2004) Gradual deformation of boolean simulations. O.Leuangthong and C.V.Deutsch(eds.), Geostatistics Banff, 939-948. Netherlands, Springer.

Liu, X., Harding, A., Gilbert, R., and Journel, A. G. (2004) A workflow for multiple-point geostatistical simulation. O.Leuangthong and C.V.Deutsch(eds.), Geostatistics Banff, 245-254. Netherlands, Springer.

Liu, X. and Journel, A. G. (2004) Improving sequential simulation with a structured path guided by information content. Math.Geology 36[No 8 Nov], 945-964.

Liu, X., Srinivasan, S., and Wong, D. (2002) Geological characterization of naturally fractured reservoirs using multiple-point geostatistics In *Proc. SPE* 75246. Tulsa, Oklahoma, SPE/DOE Improved Oil Recovery Symposium.

Lopez, S. (2003) Modélisation de réservoirs chenalisés méandriformes: approche génétique et stochastique. PhD thesis, École Nationale Supérieure des Mines de Paris.

Maharaja, A. (2004a) Hierarchical simulation of multiple facies reservoirs using multiple-point geostatistics. PhD thesis, Stanford University.

Maharaja, A. (2004b) Stochastic modeling of the rhine-meuse delta using multiple-point geostatistics. O.Leuangthong and C.V.Deutsch(eds.), Geostatistics Banff, 653-659. Netherlands, Springer.

Matheron, G. (1968) *Traité de géostatistique appliquée*. (version russe) Moscou.

Matheron, G. (1969) Structures aléatoires et géologie mathématique. Rapport Technique N-111. Centre de Géostatistique, ENSMP.

Matheron, G., Beucher, H., Fouquet, C., Galli, A., Guerillot, D., and Ravenne, C. (1987)

---

Conditional simulation of the geometry of fluviodeltaic reservoirs. In *Proc. SPE* 16753. Dallas, SPE Annual Technical Conference and Exhibition.

Okabe, H. and Blunt, M. J. (2004) Multiple-point statistics to generate pore space images. O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 763-767. Netherlands, Springer.

Remy, N. (2001) Multiple-point statistics for image post-processing. Stanford Center for Reservoir forecasting.

Remy, N. (2003) Program code SNESIM. Department of Geological and Environmental Science, Stanford University.

Roggero, F., Feraille, M., Le Ravalec-Dupin, M., and Costa Reis, L. (2005) Updating reservoir models with dynamic data and uncertainty quantification: an integrated approach. In *Proc. OAP&C Seminar*. Rueil-Malmaison, France.

Sarda, S., Jeannin, L., Basquet, R., and Bourbiaux, B. (2001) Hydraulic characterization of fractured reservoirs: simulation on discrete fracture models. In *Proc. SPE* 66398. Houston, Texas, SPE Reservoir Simulation Symposium.

Scheidt, C. and Caers, J. (2007) Using distances and kernels to parameterize spatial uncertainty for flow application. In *Proc. Petroleum Geostatistics*, Cascais, Portugal, European Association of Geoscientists & Engineers.

Strebelle, S. (2000) Sequential simulation drawing from training images. Centre for Reservoir Forecasting. PhD thesis, Department of Geological and Environmental Sciences, Stanford University.

Strebelle, S. (2002) Conditional simulation of complex geological structures using multiple-point statistics. *Math.Geology* 34[No 1], 1-22.

Strebelle, S. and Journel, A. G. (2001) Reservoir modeling using multiple-point statistics. In *Proc. SPE* 71324. New Orleans, SPE Annual Technical Conference and Exhibition.

Strebelle, S., Payrazyan, K., and Caers, J. (2002) Modeling of a deepwater turbidite reservoir conditional to seismic data using multiple-point geostatistics. In *Proc. SPE* 77425. San Antonio, Texas, SPE Annual Technical Conference and Exhibition.

Strebelle, S. and Remy, N. (2004) Post-processing of multiple-point geostatistical models to improve reproduction of training patterns. Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 979-988. Netherlands, Springer.

Strebelle, S. and Zhang, T. (2004) Non-stationary multiple-point geostatistical models. O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 235-244. Netherlands, Springer.

---

Suzuki, S., Caers, J., and Caumon, G. (2007) History matching of reservoir structure subject to prior geophysical and geological constraints. In *Proc. Petroleum Geostatistics, Cascais, Portugal, European Association of Geoscientists & Engineers*.

Suzuki, S. and Strebel, S. (2007) Real-time post-processing method to enhance multiple-point statistics simulation. In *Proc. Petroleum Geostatistics, Cascais, Portugal, European Association of Geoscientists & Engineers*.

Teles, V., de Marsily, G., and Perrier, E. (1998) Sur une nouvelle approche de modélisation de la mise en place des sédiments dans une plaine alluviale pour en représenter l'hétérogénéité. *Comptes rendus de l'Académie des Sciences*, 327., 597-606.

Zhang, T., Bombarde, S., Strebel, S., and Oatney, E. (2006) 3D porosity modeling of a carbonate reservoir using continuous multiple-point statistics simulation. SPE 96308. SPE Journal 11 N 3. Dallas, SPE Annual Technical Conference and Exhibition.

Zhang, T., McCormick, D., Hurley, N., and Signer, C. (2007) Applying multiple point geostatistics to reservoir modeling - a practical perspective. In *Proc. Petroleum Geostatistics, Cascais, Portugal, European Association of Geoscientists & Engineers*.

Zhang, T., Switzer, P., and Journel, A. G. (2004) Sequential conditional simulation using classification of local training patterns O.Leuangthong and C.V.Deutsch(eds.), *Geostatistics Banff*, 265-273. Netherlands, Springer.

## ANNEXE

---



## Annexe A. Notation

$I$	la grille de l'image d'apprentissage
$E$	la grille de l'espace de simulation
$K$	le nombre de faciès présents
$u_0$	le point courant de simulation, le point central de la fenêtre d'exploration
$\mathbf{u} = (u_1, u_2, \dots, u_n)$	l'ensemble des points voisins déjà simulés autour de point central $u_0$
$L$	le nombre maximal de points voisins autour du point $u_0$ , la taille de fenêtre d'exploration correspond au dernier niveau de l'arbre de stockage
$l$	le nombre de points voisins ( $l = 0..L$ ) pris en compte pendant la simulation du point $u_0$ correspond à niveau $l$ de l'arbre de stockage
$\mathbf{v} = (v_1, v_2, \dots, v_L)$	l'ensemble des points voisins qui définit la fenêtre d'exploration de taille $L$
$\mathbf{v}^l = (v_1, v_2, \dots, v_l); l: 0..L,$	l'ensemble des points voisins qui définit la fenêtre d'exploration de taille $l$
$Z(\cdot)$	la fonction aléatoire principale catégorielle; elle peut prendre $K$ états différents
$S(\cdot)$	la fonction auxiliaire continue, la contrainte, $S \in R$
$P[Z(u_0)]$	la loi de $Z$ au point $u_0$
$P[Z(u_0) Z(\mathbf{u})]$	la loi conditionnelle de $Z$ au point $u_0$ connaissant les valeurs de son voisinage
$P[Z(u_0) Z(\mathbf{v})]$	la loi conditionnelle de $Z$ au point $u_0$ connaissant les

---

	valeurs dans le voisinage vu dans la fenêtre d'exploration
$P[Z(u_0) Z(\mathbf{v}^l)]$	la loi conditionnelle de $Z$ au point $u_0$ connaissant les valeurs dans le voisinage vu dans la sous fenêtre de taille $l$
$P[Z(u_0) Z(\mathbf{u}), S(u_0)]$	la loi biconditionnelle de $Z$ au point $u_0$ connaissant les valeurs dans son voisinage et la contrainte $S$ au point $u_0$
$P[Z(u_0) Z(\mathbf{v}^l), S(u_0)]$	la loi biconditionnelle de $Z$ au point $u_0$ connaissant les valeurs dans le voisinage de taille $l = 0..L$ et la contrainte $S$ au point $u_0$
$S\{O(w_0)\}$	l'ensemble des points voisins qui affectent la valeur de la fonction auxiliaire ; le support de la fonction auxiliaire
$P[S > s   Z\{O(w_0)\}]$	la loi empirique de $S$ extraite directement des images d'apprentissage et conditionnée à l'état du support



## Annexe B. Caractérisation d'image

Dans cette section, nous présentons différentes méthodes statistiques et morphologiques pour comparer les géométries d'une structure simulée (par l'approche MP) et de l'image d'apprentissage.

Une simulation MP est une simulation stochastique permettant la reproduction de structures aléatoires. Une structure peut aussi être décrite par des paramètres géométriques tels que la forme, l'étendue, l'orientation etc. C'est la raison pour laquelle on tient à utiliser des caractéristiques statistiques aussi bien que morphologiques. Nous allons regrouper ces caractéristiques selon leurs ordre de complexité : 1) **Statistiques de premier ordre** (moyenne) ; 2) **Statistiques de deuxième ordre** (variance, covariance, variogramme ...). Comme la suite de cette section va la montrer, les statistiques à deux points ne sont pas capables de capturer la complexité de patterns géométriques. Aussi se tourne t-on vers les 3) **statistique multipoints** et **caractéristiques morphologiques** qui sont été souvent utilisées en traitement d'image [Baudrier (2005)].

En utilisant ces caractéristiques statistiques, deux types d'information peut être obtenus : le premier synthétise l'information de toute l'image et forme donc une statistique globale ; le second tient compte des différences au sein même de l'image et rend compte d'une information locale. Ainsi, la comparaison entre des structures géométriques pour déterminer leur degré de similarité peut prendre les deux aspects – globale et locale.

### B.1. Statistiques monovariabiles

#### *B.1.1. Moyenne globale*

Dans le cadre de ce travail, la moyenne globale est utilisée pour évaluer la proportion de faciès aussi bien dans l'image d'apprentissage que sur les réalisations obtenues par simulation MP. Dans son principe, l'approche MP est censée reproduire la proportion globale de faciès si les conditions d'application (§ 2.3) sont respectées. En pratiques, ces conditions ne sont pas souvent satisfaites ; par conséquent la proportion des faciès d'une réalisation doit être davantage contrôlée.

En outre, la précision de la reproduction des proportions dépend à la fois de la taille des

patterns géométriques vis-à-vis de la taille d'images d'apprentissage et de celle du champ de simulation. Plus l'image d'apprentissage est grande vis-à-vis des patterns, plus la proportion de faciès sera reproduite avec précision.

Pour décider si deux proportions globales de faciès sont similaires, on peut prendre une valeur  $\varepsilon$  qui spécifie le maximum acceptable entre les deux proportions. Le choix de  $\varepsilon$  est à spécifier au cas par cas. Ce choix peut dépendre de multiples paramètres tels que la taille de patterns répétitifs par rapport à la taille de l'image d'apprentissage, la taille de la fenêtre d'exploration, fiabilité de l'image d'apprentissage etc.

Exemple. Supposons que l'image d'apprentissage et l'espace de simulation soient tous deux de dimension 100x100 et la taille des patterns soit inférieure à celle de la fenêtre d'exploration qui est 10x10. Dans ce cas, les 4 conditions d'application sont respectées, l'erreur de représentation des faciès doit être lié seulement à la fluctuation statistique et à l'effet de bord. Soit  $\varepsilon = 0.01$ . Nous convenons de dire que les proportions globales  $\rho^{ref}$  et  $\rho^{MP}$  de faciès dans l'image d'apprentissage et la simulation sont similaires si  $|\rho^{ref} - \rho^{MP}| < \varepsilon$ .

### B.1.2. Moyenne locale

La moyenne locale se présente comme une carte de valeurs, de même dimension et de même taille que l'image étudiée. Ces valeurs représentent la proportion des faciès dans un certain voisinage défini par le support de valeur auxiliaire (§ 4.1.1). La précision des valeurs de proportion locale des faciès dépend fortement de la taille de fenêtre glissante. Il est important donc d'utiliser la taille de fenêtre glissante équivalente au support.

Dans les cas d'une simulation stationnaire, la moyenne locale déduite de l'image d'apprentissage peut servir de référence aux fluctuations statistiques des simulations. Si une dérive importante se manifeste sur la carte de proportion locale de faciès, il est préférable de tenir compte de cette variabilité et d'effectuer une simulation non stationnaire. Dans ce cas, cette carte de proportion locale sera utilisée comme 1) une image d'apprentissage auxiliaire et si nécessaire 2) une contrainte spatiale. Pour valider les simulations non stationnaires où le paramètre de non stationnarité est une proportion locale des faciès, nous tenons à comparer les cartes de moyenne mobile.

## B.2. Statistiques bivariées

### B.2.1. Variogramme

Souvent, pour reproduire une caractéristique géométrique d'un milieu hétérogène, on part du variogramme qui fournit des statistiques spatiales entre couples de points. On essaie ensuite de reproduire ces statistiques par une simulation stochastique d'un modèle probabiliste que l'on a spécifié au préalable. Nous pouvons donc l'utiliser le variogramme pour contrôler la reproduction de la structure géométrique en fonction de la distance. Cependant, il ne faut pas oublier que le variogramme n'est pas capable d'appréhender certaines structures géologiques complexes telles qu'un système de chenaux méandriformes. L'exemple suivant montre 3 structures différentes de variogramme similaire dans les deux directions principales (Fig. 90).

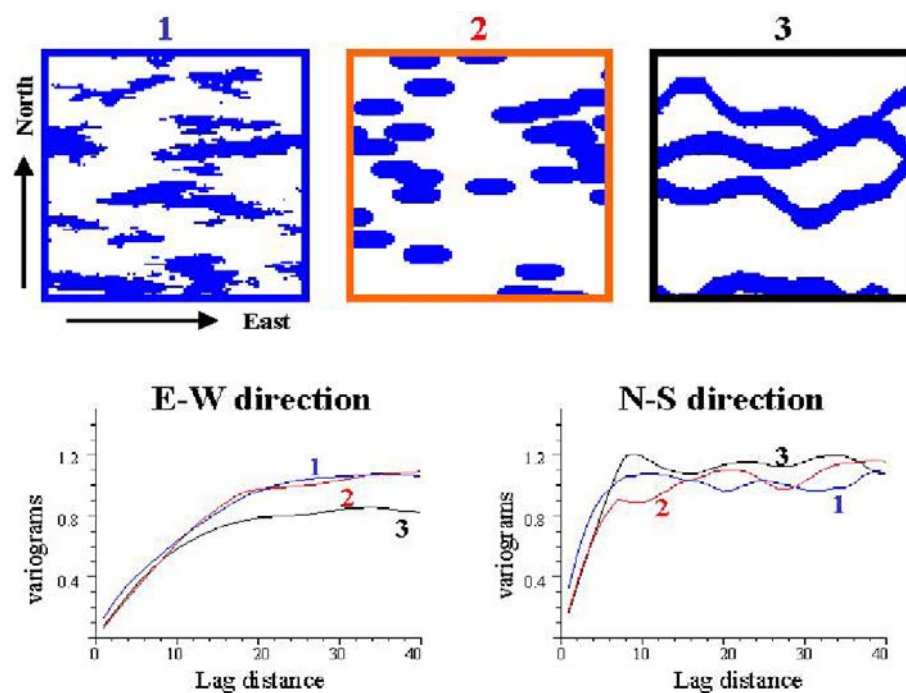


Fig. 90 Exemple de 3 structures différentes et les variogrammes qui leurs correspondent. (Liu et al., 2004)

## B.3. Statistiques multivariées

Les statistiques de deuxième ordre ne sont pas suffisamment puissantes pour saisir toutes les caractéristiques d'une structure géométrique complexe. Fondées sur des notions ensemblistes et topologiques, les statistiques multivariées et morphologiques sont plus aptes à caractériser des géométries complexes en deux et en trois dimensions.

Voici quelques méthodes de caractérisation morphologique.

### B.3.1. Connexité

Une caractéristique qui n'est pas capturée par variogramme est **la connexité** entre objets géologiques. Cette caractéristique est très importante, car elle est responsable de la circulation des fluides dans un réservoir. Les valeurs que prend la fonction de connexité renseignent sur la longueur directionnelle des clusters. Comme le variogramme, c'est une fonction d'un couple de points, caractérisé par le vecteur qui les sépare en termes de distance et d'orientation. Diverses définitions de cette fonction sont proposées dans la littérature. Dans le cadre de cette étude, on utilise la version classique proposée par Allard (1994).

Soit  $G$  l'espace discrétisé. L'appartenance d'un point  $u$  à un faciès donné  $A \subset G$  est caractérisée par sa fonction indicatrice :  $Z(u) = 1$  si  $u \in A$  et  $Z(u) = 0$  sinon.

Déf. Deux points  $u$  et  $u'$  sont proches voisins s'ils sont adjacents sur le graphe induit par l'espace discrétisé.

Déf. Deux points  $u$  et  $u'$  sont connectés :  $u \leftrightarrow u'$  dans le faciès  $A$  s'ils sont reliés par une chaîne de proches voisins  $u, u_1, \dots, u_n, u'$ , tout dans le même faciès.

Déf. La fonction de connexité entre deux points  $\alpha$  est la probabilité d'avoir ces deux points connectés :

$$\alpha(u, u') = P[u \leftrightarrow u' \mid u, u' \in A] \quad u, u' \in G.$$

Déf. L'ensemble de points  $\Gamma$  est un cluster si tous ses points sont connectés :

$$\forall u, u' \in \Gamma \Rightarrow u \leftrightarrow u' \quad u, u' \in A \subseteq G.$$

### B.3.2. Connexité multipoints

La connexité multipoints est aussi une fonction qui définit le transport entre les faciès, mais elle est encore plus sensible à la disposition des points étudiés que la connexité à deux

points. Sur une grille carrée (en 2D), la fonction  $\alpha(u_{i,j} \leftrightarrow u_{i+h,j} \leftrightarrow u_{i,j+h} \leftrightarrow u_{i+h,j+h})$  ne montre pas seulement la connexité entre ces quatre points mais donne aussi la taille approximative des clusters.

Déf. La fonction de connexité entre  $n$  points  $\alpha$  est la probabilité d'avoir tous ces  $n$  points connectés :

$$\alpha(u_1 \leftrightarrow u_2 \leftrightarrow \dots \leftrightarrow u_n) = P[\forall i \leq n : u_i \in \Gamma \mid u_i, u_j \in A] \quad u_i, u_j \in G.$$

### B.3.3. Statistiques multipoints

Dans la simulation MP, les statistiques multipoints sont utilisées pour reproduire la structure géométrique. Elles peuvent aussi servir à contrôler la qualité de cette reproduction.

Les caractéristiques étudiées ici sont encore la taille, la forme et l'orientation des patterns. Basée sur la théorie des ensembles, des treillis et la topologie, la morphologie mathématique propose les approches ensemblistes. Les images discrétisées sont étudiées via leur relation avec un ensemble fixé, appelé élément structurant. Selon la terminologie du début, un élément structurant est un voisinage prédéfini et fixé :  $Z(\mathbf{u}) = (Z(u_1) = z_1, Z(u_1) = z_2, \dots, Z(u_n) = z_n) \quad z_i \leq K$

L'analyse de répliques à l'aide de différents éléments structurants peut donner une idée des états plus fréquents et moins fréquents. La comparaison des statistiques multipoints de l'image d'apprentissage avec celles des simulations permet d'estimer les fluctuations statistiques associées à l'approche MP.



## Annexe C. Critère progressif de compatibilité

Dans cette section nous présentons les critères de compatibilité élaborer pour raffiner le choix des statistiques biconditionnelles au cours d'une simulation MPNST.

Rappelons qu'à chaque état de support figurant dans l'image d'apprentissage est associée une loi de valeurs auxiliaires :  $P[S < s \mid Z\{O(w_0)\}]$ . Cette loi empirique est utilisée pour répondre à la question : « quelle est la probabilité que telle valeur auxiliaire (telle caractéristique) corresponde à tel état de support (telle géométrie) ? ». A cet intention, seuls les états de support de fréquence non nulle sont utilisés pour la simulation de la fonction aléatoire  $Z$  au point  $u_0$ . Ceux de fréquence nulle sont supprimés.

Parmi les supports considérés, certains peuvent être plus cohérents aux contraintes spatiales (i.e. de fréquence plus importante) que d'autres. Il est alors logique de leur accorder plus d'importance. Pour ce faire, il est commode d'introduire des coefficients  $\lambda_i$ . Nous appellerons *poids*  $\hat{\lambda}_i$  ces coefficients normalisés. La loi conditionnelle est construite en s'appuyant sur les poids selon la formule :

$$P[Z(u_0) \mid Z(\mathbf{u}) = \mathbf{z}, S(u_0)] = \sum_{i=1}^N \hat{\lambda}_i P[Z(u_0) \mid Z\{O(w_0)\}^i, S(u_0)]$$

dans laquelle  $N$  désigne le nombre total d'états de supports différents figurant dans l'image d'apprentissage.

### C.1. Poids approximés par le modèle gaussien

Supposons, que la distribution empirique des valeurs auxiliaires associées à un état du support  $Z\{O(w_0)\}$  peut être extrapolée par une densité gaussienne de moyenne  $m$  et de variance  $\sigma^2$ . Dans ce cas, nous pouvons estimer la compatibilité d'un état du support aux contraintes en évaluant sa distance à la moyenne normée :

$$\varepsilon = \frac{|S(u_0) - m|}{\sigma}$$

Plus  $\varepsilon$  est proche de zéro, plus cette compatibilité est importante. L'expression des coefficients (non normés) est de la forme  $\lambda = 1/\varepsilon$ .

## C.2. Poids approximés par l'histogramme

Supposons ici que la loi des valeurs auxiliaires a une forme explicitement connue. Dans ce cas, nous pouvons opérer directement avec l'histogramme empirique  $H[Z\{O(w_0)\}, s]$  qui représente la loi des valeurs auxiliaires associées à l'état du support. Il est important de garder toujours les mêmes pas de discrétisation d'un histogramme pendant toute la simulation. Les coefficients peuvent être estimés de manière suivante :

$$\lambda = H[Z\{O(w_0)\}, s = S(u_0)]$$

## C.3. Poids approximés localement

L'approximation des coefficients par la construction d'un histogramme à partir de l'image entière peut ralentir considérablement la simulation. Pour économiser du temps calcul, nous proposons ici une construction d'histogramme partiel. Seul les valeurs proches à la contrainte donnée sont utilisées pour l'histogramme :  $H[Z\{O(w_0)\}, s], s \in [S(u_0) - \varepsilon; S(u_0) + \varepsilon]$ . Au lieu de fixer des pas d'histogramme, nous choisissons un intervalle autour de la contrainte et retirons le nombre de valeurs figurant dans cet intervalle afin d'estimer la valeur d'histogramme au point  $s = S(u_0)$ . Les coefficients sont alors donnés par la formule :

$$\lambda = H[Z\{O(w_0)\}, s = S(u_0)]$$